V. Quiroga-Fonseca

Characterization of Superselective Behavior using DNA Nanostars

Exploring the influence of **binding affinity** and **structural flexibility** on DNA nanostar superselectivity

Characterization of Superselective Behavior using DNA Nanostars

Exploring the influence of binding affinity and structural flexibility on DNA nanostar superselectivity

By

V. Quiroga-Fonseca

in partial fulfilment of the requirements for the degree of

Master of Science
in Nanobiology

at the Delft University of Technology,
to be defended publicly on May 30th, 2024.

Supervisor:	Dr. ir.	, Liedewij Laan TU Delft
Thesis committee:	Dr.	, Jos Zwanikken TU Delft
	Dr.	, Michal Shemesh TU Delft

TU Delft

Preface

"Life is a long lesson in humility" – James M. Barrie

This thesis represents the culmination of my research journey into the fascinating world of DNA nanostars and their superselective behavior. More importantly, it represents the culmination of my journey through facing hardships in academia, a lesson I am sure I will never forget.

I owe a deep debt of gratitude to my supervisors, Dr. ir. Liedewij Laan and Dr. Jos Zwanikken, whose enormous support and patience have been invaluable throughout this process, both academically and personally. I also extend my thanks to my direct supervisor, Nynke Hettema, for her valuable help and encouragement during the project, even when things looked daunting. I would also like to thank my colleagues in the Laan Lab: Marieke, Leila, Enzo, Ramon, Cristine, Sophie, Ensengul, Caspar, Esther, Frank, Renske, and Zhiheng for their invaluable comments and friendly spirit during my time in the group. Additionally, I would like to thank my third reviewer, Dr. Michal Shemesh, for her support in assessing my project.

I am grateful for the financial support provided by the Foundation Justus & Louise van Effen Excellence Scholarships, which made it possible for me to study at TU Delft. On a personal note, I would like to thank my boyfriend Noah and his family for being my home away from home, and my family in Colombia and the US for their immense support even while being away. Additionally, all my thanks to Catalina and Beatriz for their warm Colombian-tinted friendship at BN, to Patricia and Tanja for all their encouraging words and actions, and to all the friends who have supported me during this process. Without your support and help, I would not be here today.

The journey of conducting this research and finishing this degree has been filled with numerous learning experiences. These experiences have made me stronger, and I will hold them for life. Thank you for being part of this ride!

V. Quiroga-Fonseca
Delft, May 2024

Contents

Table of Contents

Preface **3**
Abstract **9**
1 Introduction **10**
1.1 Multivalent Interactions in Biology 10
1.2 Experimental Approach using DNA Origami/DNA Nanostars 11
1.3 Superselectivity: theory basics 13
1.3.1 Frenkel's theory on superselectivity: Thermodynamics of Multivalent Interactions 14
1.4 Research Framework 16
1.4.1 Research Questions and Objectives 17
2 Materials and Methods **18**
2.1.1 DNA Nanostar and DNA Receptor Hybridization 18
2.1.2 DNA electrophoresis to confirm DNA nanostar hybridization. 19
2.1.3 DNA functionalized Supported Lipid Bilayer (SLB). 20
2.2 Data acquisition using TIRF Microscopy 21
2.3 Data Analysis 21
2.3.1 Intensity values into bound fraction 21
2.3.2 Image analysis for clustering 22
2.3.3 Fitting calculations for binding fraction θfit 23
2.3.4 Mann Whitney U test for experimental comparison 24
3 Results **25**
3.1 Experimental Validation using Design A 25
3.1.1 Are DA-CL and DA statistically significantly similar? 27
3.1.2 Sensitivity analysis: Calculation of Binding Constants KA, $Kintra$, and superselectivity parameters θ, α, and <n> 27
3.2 Experimental Effects of Complementary Sequence Variations in sticky end using Design A* 29
3.3 Influence of DNA Nanostar Flexibility on Superselectivity Measurements using Design C 33
4 Discussion **39**
4.1 Interpretation of experimental results 39
4.1.1 Validation of experimental model for superselectivity using Design A 39
4.1.2 Evaluation of sequence mismatch on sticky end using Design A* 40
4.1.3 Effects of increase flexibility on DNA nanostars - Design C 42
5 Conclusion **43**
5.1 Potential refinements in experimental design for further research 45
6 Bibliography **47**
7 Appendices **52**
7.1 Experimental protocols 52
DNA Nanostar and Receptor Hybridization 52

Small Unilamellar Vesicles of DOPC 52
DNA Nanostar binding experiments – sample preparation: DNA nanostar hybridization 53
DOPC SUVs: Flow channel functionalization with lipid membrane and receptors 53
Imaging parameters 53
Appendix: Dilutions for receptor densities 54
7.2 TIRF DNA Nanostar Visualization 55
7.3 $KA, Kintra$ error propagation calculations from general error propagation formula 56
7.4 Analysis Codes 56

Abstract

Multivalent interactions are crucial mechanisms employed by cells to respond to their environment, often leading to superselectivity phenomena. Extensive experimental and theoretical efforts have been made to understand the variables controlling superselectivity, but challenges persist in achieving precise control over receptor and ligand numbers in biological systems. To address this, the Laan lab developed a model system using DNA origami, enabling precise manipulation of receptor and ligand numbers at the nanoscale. These nanoscopic (~15 nm) structures can mimic the receptors of a target surface, and the extracellular ligands are represented by a branched star-shaped DNA origami structure in solution. Both structures hold a fluorophore, allowing visualization using total internal reflection microscopy (TIRF) by measuring intensity values of the DNA nanostars absorbed into target surface. In this study, we investigated the effects of altering binding strength and flexibility in DNA nanostar structures on superselectivity. We found that experimental results for replicating previous experiments using the same DNA nanostars (Design A) exhibited minor variations within expected ranges, validating the reliability of the experimental protocol. However, sensitivity analysis highlighted the influence of data points on superselectivity interpretation, emphasizing the need for careful data analysis. Our study also evaluated the impact of introducing sequence mismatches on binding affinity (Design A*), revealing that modifications reducing binding affinity do not necessarily enhance superselective behavior for this system. Additionally, our investigation into the effects of increased flexibility (Design C) revealed unexpected behaviours in bound fraction and cluster formation, suggesting a potential relationship between cluster formation, intensity values registered, and the flexibility of the structure mediated by phase separation. These findings underscore the complexity of DNA nanostar behavior and stress the need for further research to elucidate the factors influencing superselectivity in DNA nanostars. By providing more understanding into how to develop highly selective particles, therapeutic molecules could sharply distinguish between healthy and corrupt cells, leading to customizable treatments with higher efficacy.

Keywords: superselectivity, DNA nanostars, adsorption, multivalent interactions.

1 Introduction

1.1 Multivalent Interactions in Biology

For over 30 years, research focused on using nanoparticles as delivery vehicles for therapeutic molecules elucidated a world of possible applications in targeted ways using multivalent principles (Mammen et al., 1998). In this context, let us use the definition of a multivalent interaction as the specific interplay between receptors and a ligand in two different biological entities, involving multiple interactions simultaneously (Mammen et al., 1998). In general terms, receptors denote surface “receivers” on cells, while ligands refer to molecules that alter the receptor function (Figure 1). By convention, the receptor becomes the host, and the ligand becomes the guest (Mahon & Fulton, 2014). Multivalent systems typically comprise k ligands, usually linked to a "core." The number of receptors n_R and k ligands can form up to *i* bonds, where *i* is the smallest of n_R or k, considering single ligand binding per receptor site.

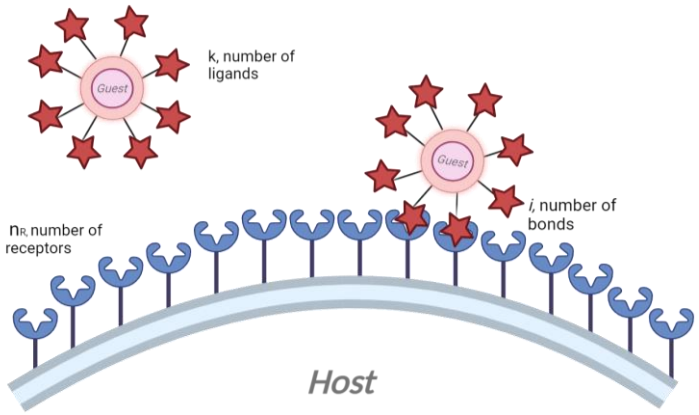


Figure 1. Schematic illustration depicting a multivalent particle (guest) interacting with a host cell covered in compatible receptors. The diagram illustrates a host cell with receptors (n_R) and a guest particle with ligands (k). Each receptor can bind with multiple ligands, forming up to i bonds, where i is the smaller of the total number of receptors (n_R) or ligands (k), considering single ligand binding per receptor site. Created with BioRender.com

Drawing inspiration from the dynamic world of how viruses, bacteria, and cells interact, the synthesis of synthetic multivalent constructs has surged a way to control biological interactions (Lombardo et al., 2019). Whiteside et al. have reported that protein interactions, virus-host dynamics, cell-cell adhesion, protein polarization and interactions, antibody action mechanisms, and other biological processes (Arsiwala et al., 2019; Tong et al., 2009) constitute some of the examples where multivalent interactions determine functionality. Given its appearance in biological contexts, multivalency has offered a possibility for various applications where ligand-receptor interactions can play a pivotal role. However, the mechanisms underlying these interactions are not easily employed experimentally (Dubacheva et al., 2015), hence the importance of further exploring experimental models with a higher degree of control over multivalent interactions is required.

Nonetheless, given the relevance of multivalent interactions in nature, this concept has radiated across diverse biological and synthetic experimental applications spanning vaccines, antibodies, carbohydrates, liposomes, dendrimers, polymers, peptides, oligonucleotides, and supramolecular constructs to improve biological applications (Arsiwala et al., 2019; Carlson et al., 2007; Lombardo et al., 2019; Mahon & Fulton, 2014; Soliman et al., 2011; Wang et al., 2020). For instance, an interesting connection between multivalency and immune response in humans has been explored by Kiessling and collaborators, by proposing some mechanisms by which multivalent ligands engage with oligomeric receptors on cell surfaces. By using the receptor clustering phenomenon, they found that multivalent ligand density controls the clustering of receptors, critical for subsequent signaling cascades that regulate immune response in humans. Another big example of applications has been around targeting cancerous cells, by using the subtle variations in receptor concentrations that cancerous cells show, to create superselective therapies or drug delivery vehicles that don’t attack healthy cells (Antignani et al., 2020; Kapcan et al., 2023; Wang et al., 2020; Zhang

et al., 2014). More recently, other applications have taken the properties of multivalent binding using DNA-origami to enhance targeting *in vivo* tumor cells (Heuer-Jungemann et al., 2016; Zhang et al., 2014), to enhance target delivery by using DNA self-assembly properties with scaffold design (Li et al., 2016), and to customize chemotherapeutic drug release, or tumor necrosis triggers (Ijäs et al., 2021). Other examples include the addition of nucleic acid aptamers to the surface of genome-free viral capsid carrier to increase its binding capabilities (Ozalp et al., 2011; Tong et al., 2009), using DNA-origami nano agents to trigger over 100x times faster rates of apoptosis in cells (Berger et al., 2021).

Although we can see the diverse and wide application space that multivalent interactions have brought into biological settings, there is still a lot of work needed to scale up these discoveries into generalized procedures for targeted medical treatments or therapeutics, as some off-target effects and poor delivery efficiency (Antignani et al., 2020; Wilhelm et al., 2016) have been the encountered challenges to obtain fully clinically approved procedures. As research of multivalency and biological landscapes grows, the lack of reproducibility in the experiments reported and the specificity of the studies conducted (sometimes using materials that are not suitable for biological applications) seems to call for adopting new hypotheses or protocols that can lead to more reliable results (Leroux, 2017; Tjandra & Thordarson, 2019; Woythe et al., 2021). This could indicate there is a gap in either the theoretical aspects of such interactions, the standardization of the experimental methods that are being used, or the utilization of computational models that support the design of such molecules. However, the need is pressing, as increasing adsorption and reducing unwanted side effects of current therapeutics seems to be limited by the lack of specificity and selectivity from drug molecules in use (Woythe et al., 2021).

Numerous medical fields such as tissue engineering, cardiology, vaccinology, and oncology have identified the benefits of targeted molecules (Woythe et al., 2021). Tuning in and out superselective behavior in particles could represent a breakthrough in systemic drug delivery therapeutics, by providing particles capable of sharply distinguishing the receptor cell where the therapeutic molecule should be delivered. Maximizing the selectivity of multivalent drug delivery systems centers on understanding the underlying factors governing multivalent ligand binding of multivalent theory, which is still a work in progress.

Therapeutic molecules must prevail over several biological barriers before arriving at their target receptors in the desired cell, then superselective behavior would ensure higher efficacy. Theoretical models show superselective interactions are highly sensitive to the number of receptors and ligands available (Linne et al., 2021; Tjandra & Thordarson, 2019), ergo having quantitative experimental approaches for this end is essential to achieve real applications. Moreover, most of the experimental models mentioned before containing ligand-receptors interactions can’t be individualized for control on low number valencies. For this end, The Laan Lab developed a simple experimental model, assembling the most important characteristics of a multivalent superselective interaction while allowing high control over ligand and receptor number in low regimes. This approach allows to measure changes in binding dynamics given by small variations in the number of receptors, ligands, and binding strength. To do so, star-shaped DNA origami structures known as DNA nanostars are used, as a molecular representation of the ligand-receptor interactions in cells. This thesis aims to extend the knowledge of superselective behavior in the nanoscale, providing an exploration of experimental reproducibility and the consideration of additional variables like flexibility, utilizing this unique experimental approach developed at the Laan Lab.

1.2 Experimental Approach using DNA Origami/DNA Nanostars

DNA nanostars are DNA origami structures composed of several single-stranded DNA (ssDNA) molecules, that result in a self-assembled DNA star-shaped structure when hybridized together (Biffi et al., 2013). These nanoscopic particles then portray a controlled number of interacting terminations, mimicking a particle with a well-defined valency. It is possible to add a "sticky end" to each DNA arm (a non-paired ssDNA sequence), complementary to the sequence that will resemble the receptor structure. The number of base pairs and specificity of the complementary sequence for the sticky end will control the interaction strength. The sequence used for the non-sticky section of the arm determines the flexibility and material properties of the DNA nanostar. Modifications on the ssDNA sequences

with fluorophores or cholesterol molecules are possible, facilitating visualization and manipulation of the experimental model. For further clarification, see figure 2.

For this thesis, we have used an experimental model to simulate superselectivity on a surface that resembles a lipid membrane. For selected sticky ends in the DNA nanostars in solution, we will embed a double-stranded DNA (dsDNA) backbone sequence to a coverslip shielded with a Supported Lipid Bilayer (SLB) composed of small unilamellar vesicles (SUVs) made from the phospholipids 1,2-dioleoyl-sn-glycero-3-phosphocholine (DOPC). We will truncate the receptors using a cholesterol molecule Chol- TGE at the 5' end, and a Cy3 fluorophore at the 3' end. For the DNA nanostars in solution, we will use different designs that can bind complementary to the selected receptors and an Atto488 fluorophore molecule will be added at the 3' end of one sticky end. A complete experimental setup is referred to in figure 2.

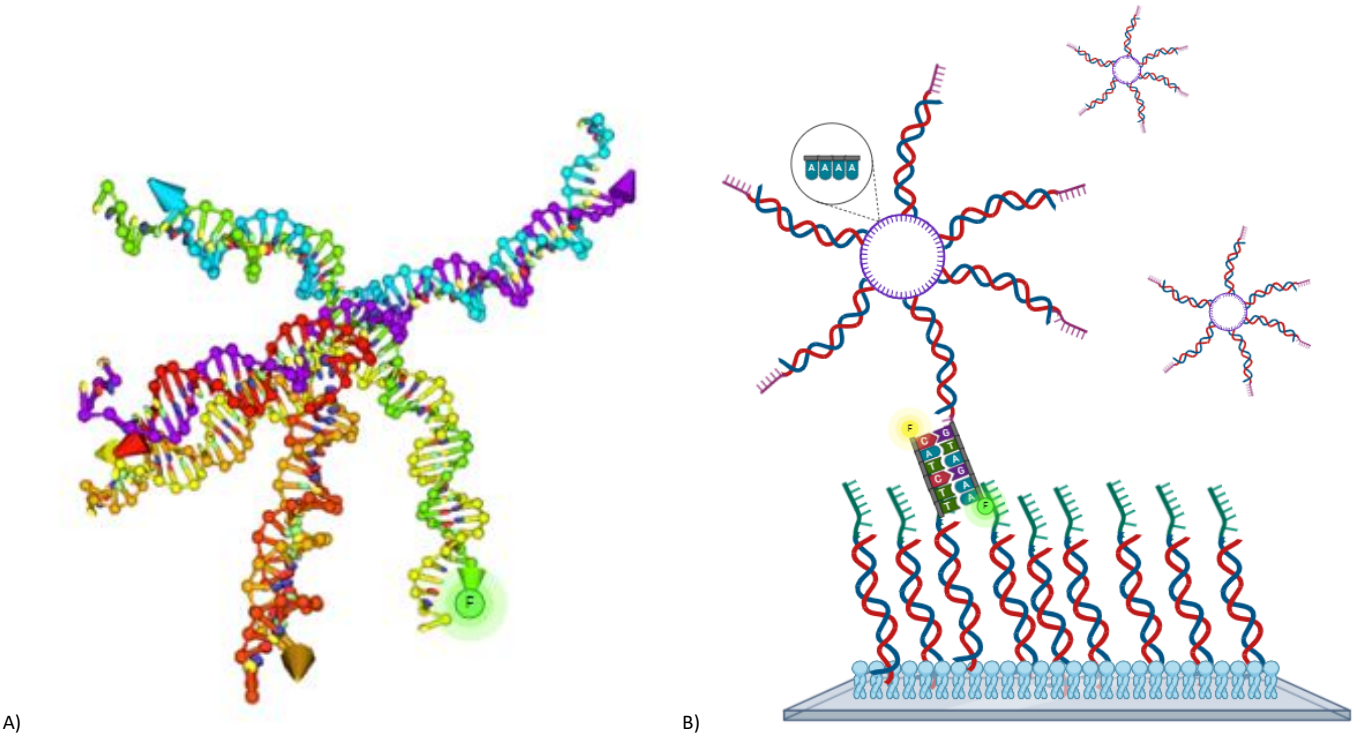


Figure 2 DNA nanostars and experimental model. A) Simulation model of a 6-arm DNA nanostar in solution. Each arm terminates with sticky ends represented by non-paired nucleotides. The Atto488 fluorophore is depicted in green at the end of one of the sticky end sequences. The central region displays non-binding base pairs, typically composed of repetitive oligo sequences such as poly-A or poly-T. This simulation was created using oxView (Poppleton et al., 2021). B) Experimental design illustrating a target surface containing receptors, along with DNA nanostars in solution to mimic ligands. Both the receptors and DNA nanostars are labeled with fluorophores to aid visualization using Total Internal Reflection Fluorescence (TIRF) microscopy. The blue section denotes the Supported Lipid Bilayer (SLB) where receptor sequences are truncated using cholesterol-modified DNA sequences to a microscopy slide. The green and yellow dots denote the corresponding Atto488 and Cy3 fluorophores in the DNA nanostar and receptor sticky ends, respectively. Not to scale. Created with BioRender.com

The addition of fluorescent molecules will allow the direct visualization of the nanostar surface adsorption using total internal reflection (TIRF) microscopy. By measuring light intensity levels emitted by the sample after being exposed to a 488nm excitation laser, we will be able to quantify the number of nanostars attached to the receptors on the surface. By varying the sticky end length, the valency of nanostars, and the base pairs that conform to the nanostar's center, it is expected to detect changes in the average intensity values registered. A brief explanation of TIRF microscopy is depicted in Figure 3.

For our experiments, we used only k=6 DNA nanostars, with modifications in vertex sequence and base pair number (Design C), mismatch in the base pair number between arm sticky end and receptor sticky end (Design A*), and the unmodified design used by Linne (2022) (Design A).

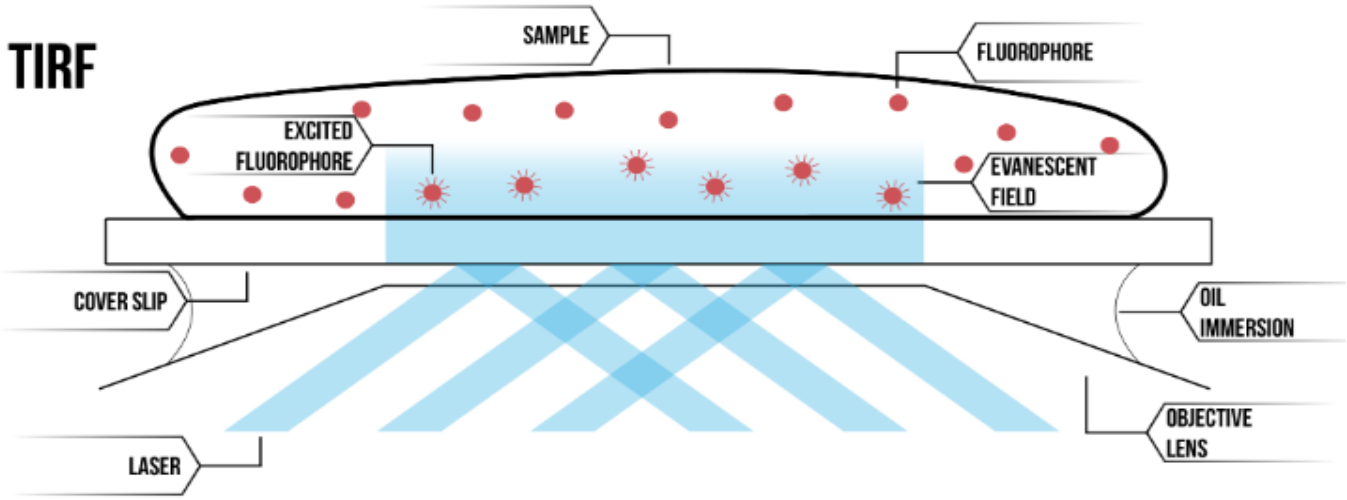


Figure 3 Schematic of TIRF microscopy mechanism of action. By irradiating the sample with a low-power laser (typically 15% of the laser wavelength) within the absorption spectra of the fluorophores, an evanescent wave is generated within the coverslip. This wave causes excitation of fluorescent molecules near the surface, resulting in emission at a lower wavelength. This enables clear separation of the emitted light from the incident laser light at the collecting camera. Adapted from (Capsid Constructors, n.d.).

1.3 Superselectivity: theory basics

Measuring superselectivity is the core of our experimental approach, and is defined as the ability of multivalent particles to distinguish and select surfaces to which it will bind, based on the accessible receptors at a target surface (Dubacheva et al., 2015; Linne, 2022), as shown in Figure 4. The ligand particles can detect small changes in the number of receptors, making superselectivity a highly precise process. Most of the theoretical frameworks we have about superselectivity in multivalent interactions rely on the works of Whitesides and collaborators (Mammen et al., 1998); and Frenkel and collaborators (Dubacheva et al., 2015; Martinez-Veracoechea & Frenkel, 2011, p.) where they have established the variable convention and in-depth exploration of multivalent binding thermodynamics. However, experimental measurements to corroborate the theory in biological cases represent a technical challenge since the chances of detecting the effects of individual physical properties in such nanoscales are limited (Linne, 2022). To this end, let us dive into the basics of superselectivity, and some examples of experimental protocols developed to study it.

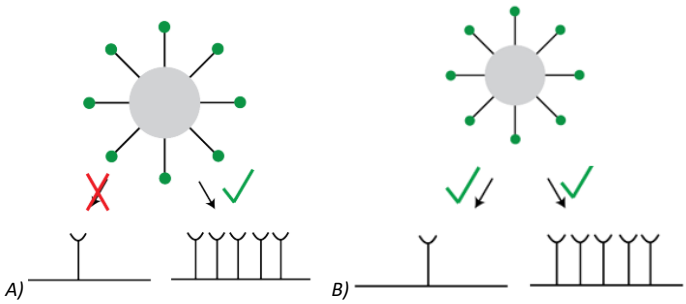


Figure 4 A) Schematic figure of superselectivity, the ability to select surfaces based on the number of receptors. B) Schematic figure of multivalency, the ability to bind to one or multiple receptors. Adapted from Linne (2022).

1.3.1 Frenkel's theory on superselectivity: Thermodynamics of Multivalent Interactions

This section explores the interplay of thermodynamics and kinetics underlying multivalent interactions, showing how these physical principles can dictate the success of binding selectivity in experimental applications. As mentioned before, the pioneering works of Frenkel and Whitesides to standardize multivalent binding thermodynamics are at the moment widely used in the field but come with certain limitations by often ignoring competition from neighboring binders, steric forces on receptors, kinetic barriers due to clustering, and local structural complexities near the cell membrane (Curk et al., 2018), factors that might be more important for superselective behavior than previously considered.

A great part of the multivalent theory up to now heavily relies upon statistical thermodynamics, explaining multivalency as a statistical phenomenon bridging physics and biology. Explaining selectivity can be done by using two equations: adsorption profile Θ, and the selectivity factor α. A critical, and well-defined, receptor density allows for "switch-like" adsorption dynamics (Figure 5), meaning that the interaction can turn on and off. Such a sharp response is then defined as superselectivity. Thus, the adsorption profile, or bound fraction, can be written as (Curk et al., 2018; Martinez-Veracoechea & Frenkel, 2011):

$$\theta = \frac{\rho_n * K_A^{av}}{1 + \rho_n * K_A^{av}} \tag{1.1}$$

where ρ_n is the molar concentration of DNA nanostars in solution, K_A is the affinity constant that describes the bonding affinity for an individual ligand-receptor binding event, and K_A^{av} is the equilibrium avidity association constant, which is the accumulated strength of all the multiple affinities. K_A^{av} is also described in terms of a new variable, K_{intra} a constant representing the likelihood of forming subsequent bonds after the first bond was formed, and whenever cooperative effects among the ligands are ignored. Defining σ_R as the density of available receptors on the host, the relation among K_A and K_{intra} describes the formation of all possible bonds as (Curk et al., 2018)(Figure 5):

$$K_A^{av} \sim \frac{K_A}{K_{intra}} \left[(1 + n_R K_{intra})^k - 1\right] \tag{1.2}$$

with n_R and k representing the total number of receptors and ligands. For our experiments, we will use an adaptation of these term employed by Linne in her experiments, where the space of ligands and receptors is enclosed:

$$K_A^{av} \sim \frac{K_A}{K_{intra}} \left[(1 + \sigma_R A K_{intra})^k - 1\right] \tag{1.3}$$

where $\sigma_R A$ represents the number of finite receptors available for binding within a total surface area A. Since θ depends on K_A^{av}, this expression shows that the bound fraction is sensitive to the number of perceived receptors by the guest on the host, a pivotal factor in targeted binding scenarios. Let us note that both equations assume there are no effective interactions (e.g. attractions or repulsions) among DNA nanostars. Now, to quantify the selectivity factor α, we use the definition (Martinez-Veracoechea & Frenkel, 2011):

$$\alpha = \frac{d\ln \theta}{d\ln \sigma_R} \tag{1.4}$$

for which a sharp switch-like behavior concerning the number of receptors will be reflected in a heightened selectivity. In other words, a system is called **superselective** (Figure 5), if the relation between the receptor concentration and the binding probability is nonlinear, such that:

$$\theta = (\sigma_R)^{\alpha}, \alpha > 1 \tag{1.5}$$

These equations also unravel how multivalent selectivity centers upon various factors, including binding strength (K_{intra}), ligand number (k), and receptor availability (σ_R). Since our experimental model allows for control over low valence numbers, the next question that arises is how many ligands are effectively binding to the receptors. For this, we will use a more specific analytical expression (Linne, 2022) developed to calculate the average number of bound arms after the first bond was formed, using DNA nanostars:

$$< n >= 1 + \frac{k-1}{(1+\sigma_R A K_{intra})^{-1}} \tag{1.6}$$

Let us note that Eq. 1.6 relies on the assumption that the superselective behavior of DNA nanostars only depends on the receptor density and the likelihood of establishing new bonds after the primary bond. The equations above established our theoretical framework to study superselectivity in DNA nanostars. In the next section, we will show how these are related to our experimental results.

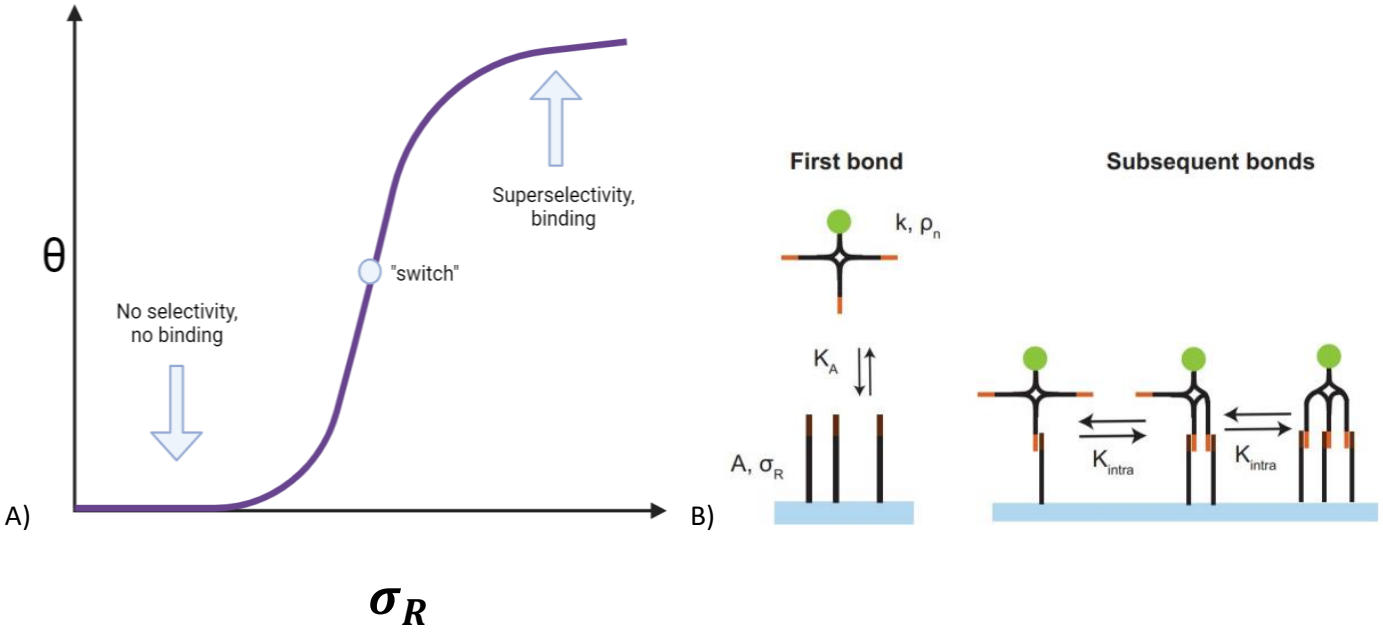


Figure 5 Theoretical "Switch-Like" Relation Bound fraction. A) This schematic illustration demonstrates the theoretical "switch-like" relationship between the adsorption profile, receptor density, and superselective behavior. The graph depicts the non-linear transition of the bound fraction θ, showcasing how selectivity shifts from non-selective to superselective as receptor density increases. Created using BioRender.com. B) Schematic illustration showing difference between K_A and K_{intra}. Adapted from (Linne, 2022).

1.4 Research Framework

Linne (2022) was able to provide experimental evidence for the binding dynamics of DNA nanostars with 1,3,6 and 10 arms on a target surface allowing the testing of Frenkel's theories about superselectivity and valency number correlation. Frenkel stated that the bound fraction of the guest particles (DNA nanostars in this case) on the surface should sharply increase nonlinearly with the number of arms it holds (Dubacheva et al., 2015). However, Linne (2022) found larger superselectivity values for DNA nanostars with fewer arms, where DNA nanostars with valency k= 3 arms showed the most superselective behavior over k=6 or k=10 (Figure 6). Even though this optimum has been quantitatively explained by proposing extending the current theory with interactions between DNA nanostar binding sites (Linne et al., 2023), we also would like to propose this could correspond to a technical limitation, as the considerations for the model state that: i) there should not be cooperation among receptors and ligands, and ii) that the DNA nanostars should do not interact among each other, meaning the binding should correspond to individual DNA nanostars interacting multivalently with the surface. Moreover, if we reflect DNA nanostars from a structural point of view, the flexibility of each arm and how structurally mobile they are could potentially determine the number of arms that are effectively binding to the target surface. Thus, we think the configuration of the arms and their flexibility in the DNA nanostars could be a limiting factor for the superselective behavior registered with our experimental method.

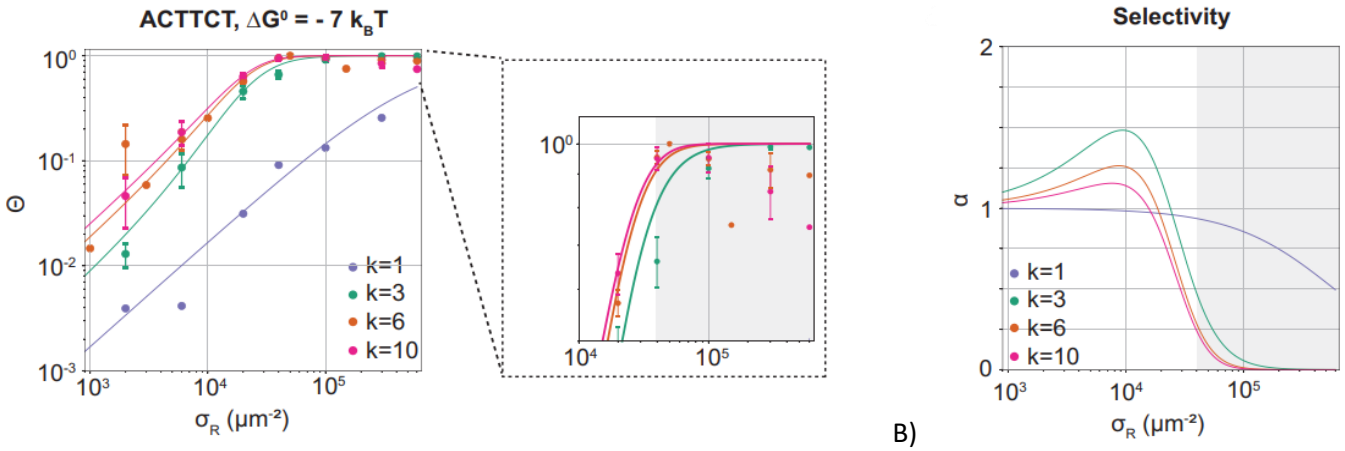


Figure 6 Results from Linne (2022) A) Measurements of bound fraction vs. receptor density values on the target surface. Experiments were conducted separately for DNA nanostars of 1,3,6 and 10 arms. The grey area indicates regions where experimental data deviates from the theoretical model, revealing a reduction in bound fraction for nanostars with higher valency numbers. B) Selectivity calculations from bounded fractions. The critical value for superselectivity corresponds to the 3-arm nanostar. Adapted from (Linne, 2022).

Further investigation into the dynamic properties of DNA has uncovered interesting findings that could provide insight into the limitations of the current experimental protocol. Specifically, the persistence length of DNA differs significantly between single-stranded DNA (ssDNA) and double-stranded DNA (dsDNA), with dsDNA measuring approximately 50 nm (equivalent to around 150 base pairs) (Marin-Gonzalez et al., 2021) and ssDNA measuring about 0.75 nm (Saran et al., 2020). In our experimental setup, this difference seems acceptable, considering that each receptor is 77 base pairs long, while the arms of each DNA nanostar consist of 53 base pairs. However, studies suggest that the dynamics at the base pair level are heavily influenced by the DNA motif or sequence present in the sequence, potentially affecting the flexibility and persistence length of the DNA (Marin-Gonzalez et al., 2021). For example, research has shown that Poly A sequences are more rigid at a local level compared to Poly T sequences (Saran et al., 2020), which could be impacting experiments involving DNA nanostars shown by Linne (2022), as all the DNA nanostar designs used hold poly A sequences in the vertex. Additionally, the interaction between these DNA processes and hybridization has been found to facilitate various physical phenomena that may be relevant to our experimental setup, such as precipitation or liquid-liquid phase separation (King & Shakya, 2021).

As an example, researchers such as the Rogers Lab in Brandeis University, are currently using DNA nanostars as the proposed system to study phase separation and the complex structures formed by them below ~45°C , which is called the upper critical solution temperature (UCST), where they claim the DNA nanostars are in the two-phase

region (Hegde et al., 2023). The research shows that nanostar droplets can spontaneously form hyperuniform structures, a type of disordered material with “hidden order” that combines the long-range order of crystals with the short-range isotropy of liquids, which potentially could change the binding dynamics we are considering with our model. Since the experimental protocol we proposed does not regulate temperature, and is always lower than the given UCST, we also propose that this physical phenomenon could be affecting the superselectivity measurements obtained by Linne (2022), and our experiments as well.

1.4.1 Research Questions and Objectives

For this project, it is our interest to further explore the reproducibility of the experimental model using DNA nanostars to study superselective events (Design A), to evaluate the effects of changing the binding strength between DNA nanostars and receptors by creating a slight mismatch in the complementary sequences of the sticky ends (Design A*), and to reconsider the influence of flexibility changes in the DNA nanostar structure (Design C), by studying how it may be influencing the reported results of bound fraction Θ, selectivity factor α, average bound arms <n>, and the binding affinity constants K_A and K_{intra}. Then our research question states:

How do experimental validation, the exploration of variations in the binding affinity constants, and flexibility constraints governing DNA nanostars contribute to our understanding of superselective behavior for k=6 DNA Nanostars? What insights do these investigations provide into the deviation from theoretical predictions, particularly Frenkel's model?

And the derived sub-questions:

1. **Experimental Validation (Design A)**
    a. **Verification of Reproducibility:** How well do the replicated experimental results align with the outcomes obtained in the original studies?
    b. **Consistency Across Different Experimental Conditions:** To what extent does the superselective behavior remain consistent under varying experimental conditions, confirming the reliability of the developed technique?

2. **Effect of Complementary Sequence Variations in the sticky end (Design A*)**
    a. How do changes in the number of base pairs (bps) in the overhang sticky end complementary sequences influence the superselective behavior of k=6, 6bps sticky end DNA Nanostars?
    b. How do these differences change the effective binding strength constants?

3. **Influence of DNA Nanostar Flexibility on Superselectivity (Design C)**
    a. Image Analysis: How can the analysis of intensity values derived from TIRF microscopy images be employed to reinterpret the experimental measurements, as the values of coverage, bonding, and superselectivity derive from this crucial step?
    b. **Role of Polynucleotide Sequence in Superselectivity:** What observable effects, if any, does the variation in the polynucleotide sequence have on the measured superselective behavior?
    c. **Comparison with Theoretical Predictions:** How do the observed results align with Frenkel's model, and what insights can be gained regarding the influence of flexibility constraints on superselectivity?

To this end, we only focused on 6-arm nanostars by using the same designs for nanostars and receptors that Linne (2022) used; and added a new design where the flexibility of the arms along the structure was enhanced by altering the polynucleotide sequence along the vertex (see Figure 2).

2 Materials and Methods

2.1.1 DNA Nanostar and DNA Receptor Hybridization

For our experiments, we used only k=6 DNA nanostars, with modifications in vertex sequence and base pair number (Design C), mismatch in the base pair number between arm sticky end and receptor sticky end (Design A*), and the unmodified design obtained from previous work conducted by Linne (2022) (Design A). The sticky end length dictates the hybridization energy (ΔG = −7 k_BT). Previous experiments with colloidal particles reported by Linne (2021) indicated that superselective binding occurs around such enthalpies, which is why these sticky ends were chosen. The experimental protocols used are based on Linne (2022) work, with alterations done for the purposes of this thesis.

Each DNA nanostar is formed by combining modified six ssDNA sequences in equal concentrations (X1 to X6), where one (X1) of the strands holds a modification of a Atto488 fluorophore (Table 1). Receptors (Rec) in use are composed of two parts, receptor backbone (R1), holding the cholesterol modification, and the receptor fluorophore (R2), holding the complementary sticky end sequence that binds with the DNA nanostars, and a Cy3 modification (Table 2). All modified DNA strands were purchased from Integrated DNA Technologies Inc (IDT), resuspended in Tris Acetate-EDTA-NaCl (TAE,100 mM NaCl, pH=8) to a final stock concentration of 200 nM and stored at −20C.

To hybridize the DNA nanostars and receptors, the ssDNA were thawed and added in equal concentrations until the mixes reached a final volume of 30 µl, in a final concentration of 0.5 μM. For the annealing process, the final mixes went in a PCR thermocycler overnight, beginning in 95◦C for 10 min and progressively cooling down at a rate of 0.2◦C/min until reaching 4◦C. The hybridized DNA was later stored at 4C until used for the experiment on the same day. For the experiments, we diluted the DNA nanostars to a final concentration of 0.01uM and receptors in 6 different concentrations in Tris Acetate-EDTA-NaCl (TAE ,100 mM NaCl, pH=8, 10 mM magnesium chloride (MgCl)). We used the same buffer batch in all experiments.

Name	Sequence (5' to 3')	5'	3'
DA/DA* - X1	CTACTATGGCGGGTGATAAAAAAACGGGAAGAGCATGCCCATCCA-sticky end	-	ATTO488
DA/DA* - X2	GGATGGGCATGCTCTTCCCGAAAACTCAACTGCCTGGTGATACGA-sticky end	-	-
DA/DA* - X3	CGTATCACCAGGCAGTTGAGAAAACATGCGAGGGTCCAATACCGA-sticky end	-	-
DA/DA* - X4	CGGTATTGGACCCTCGCATGAAAACCATGCTGGACTCAACTGACA-sticky end	-	-
DA/DA* - X5	GTCAGTTGAGTCCAGCATGGAAAACGCATCAGTTGCGGCGCCGCA-sticky end	-	-
DA/DA* - X6	GCGGCGCCGCAACTGATGCGAAAATTTATCACCCGCCATAGTAGA-sticky end	-	-
DC - X1	CTACTATGGCGGGTGATAAATTTTTTTTTTCGGGAAGAGCATGCCCATCCA-sticky end	-	ATTO488
DC - X2	GGATGGGCATGCTCTTCCCGTTTTTTTTTTCTCAACTGCCTGGTGATACGA-sticky end	-	-
DC - X3	CGTATCACCAGGCAGTTGAGTTTTTTTTTTCATGCGAGGGTCCAATACCGA-sticky end	-	-
DC - X4	CGGTATTGGACCCTCGCATGTTTTTTTTTTCCATGCTGGACTCAACTGACA-sticky end	-	-
DC - X5	GTCAGTTGAGTCCAGCATGGTTTTTTTTTTCGCATCAGTTGCGGCGCCGCA-sticky end	-	-
DC - X6	GCGGCGCCGCAACTGATGCGTTTTTTTTTTTTTATCACCCGCCATAGTAGA-sticky end	-	-
DA/DC - X sticky end	GTAGAA	-	-
DA* - X sticky end	GTAG	-	-

Table 1. DNA sequences of nanostars used in the study, categorized by design, where Design A is DA, Design C is DC, and Design A is DA). The DNA strands labeled with 'X' contribute to the formation of the DNA nanostars, with the number of strands and sticky ends determining the valency of each nanostar. The strands indicated in the name were specific to the corresponding design (DA, DC, or DA*), while those without clarification were used for all three designs.*

Name	Sequence (5' to 3')	5'	3'
Receptor sticky end	TTCTAC	-	-
Receptor backbone	TCGTAAGGCAGGGCTCTCTAGACAGGGCTCTCTGAATGTGACTGTGCGAAGGTGACTGTGCGAAGGGTAGCGATTTT	Cholesterol-TEG	
Receptor	TTTATCGCTACCCTTCGCACAGTCACCTTCGCACAGTCACATTCAGAGAGCCCTGTCTAGAGAGCCCTGCCTTACGA-sticky end	Cholesterol-TEG	Cy3

Table 2 DNA receptor sequences employed in the study. Each receptor consists of a backbone sequence hybridized with specific receptor sequences, resulting in the formation of receptors with sticky ends. These receptors are designed to bind complementary to the DNA nanostars sticky end.

2.1.2 DNA electrophoresis to confirm DNA nanostar hybridization.

For hybridization verification, we performed DNA electrophoresis for each of the ssDNAs that conform the hybridization of nanostars and receptors, and their respective mixes after the hybridization protocol. For each we loaded 10 µL of 0.5 µM, on a 1.5% agarose gel, comparing against Quick-Load Low Molecular weight DNA ladder from Biolabs. After 45 min at 120 V, we obtained the following figures for Design A and Design C, and Receptors (from left to right, Figure 7.). The highest bands on each case correspond to the formation of the hybridized product, compared to the size of each individual ssDNA. Since these bands are the brightest, it indicates that most of the sample hybridized correctly. The lower bands could correspond to incomplete hybridization products as ssDNAs can have structural changes that impede complete formation of the DNA nanostar.

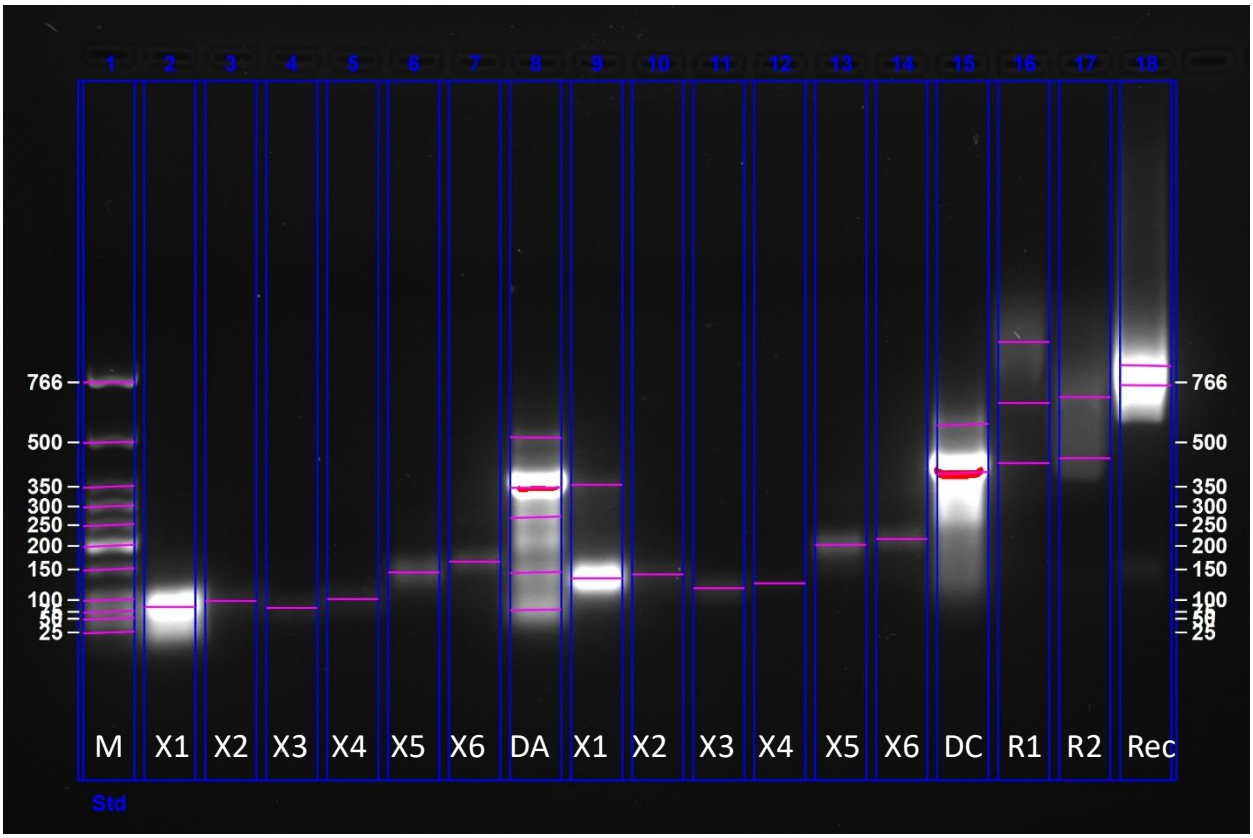


Figure 7. Gel electrophoresis results showing DNA nanostars and receptors. Lane 1: Low molecular weight DNA ladder (M). Lanes 2 to 8: Design A (DA) DNA nanostars. Lanes 9 to 15: Design C (DC) DNA nanostars. Lanes 16 to 18: Receptors (Rec). The largest band in each lane corresponds to the hybridization product. Design A DNA nanostars appear to be approximately 350 base pairs (bps), Design C DNA nanostars approximately

*400 bps, and receptors approximately 766 bps in size. The gel electrophoresis results provide a relative indication of the size difference between individual DNA strands and the hybridized DNA nanostars. Due to interference from the Chol-TEG cholesterol molecule and the three-dimensional structure of the DNA nanostars during the electrophoresis process, precise size determination is challenging. However, the focus is on assessing the relative difference between individual strands and the hybridized DNA nanostars, which is sufficient for the intended analysis. Imaged processed with Image Lab (*Image Lab Software | Bio-Rad*, n.d.).*

2.1.3 DNA functionalized Supported Lipid Bilayer (SLB).

To assess DNA nanostar adsorption on a surface covered with complementary receptors, we covered a glass microscopy slide with a supported lipid bilayer (SLB) to which the cholesterol molecules could truncate to, maintaining the DNA receptors on surface. We created small unilamellar vesicles (SUVs) using 18:1 1,2-dioleoyl-sn-glycero-3-phosphocholine (DOPC) lipids in stock. The latter were obtained from Avanti Polar Lipids and stored in chloroform. The DOCP lipids were dried out in an overnight vacuum desiccator and stored at -20C until use.

To generate the SUVs, the dried lipids were resuspended in TAE-NaCl buffer and low speed shaker for 30 mins without temperature regulation (room temperature). The obtained solution was extruded using an Avanti mini extruder, with a membrane 30 nm pore size, also from Avanti Polar Lipids. These is done to regulate the size of the vesicles obtained.

To prepare the glassware for attaching the SUVs and ensure uniform coverage, we subjected the glass slides and coverslips from VRW to an ultra-cleaning process. Initially, the glassware underwent three phases of sonication. First, they were sonicated in a 2% Hellmanex solution, followed by acetone (>99.9%), and finally in a potassium hydroxide solution (KOH, 1 M, Merck), with each phase lasting 30 minutes. Cleaning with MiliQ water was performed in between each step, ensuring thorough rinsing. All liquids were poured until they completely covered the glass slides in a 250mL beaker. For the final batch of experiments (Design C), only the glass slides were cleaned using a 2% Hellmanex solution and MiliQ water, while the protocol for the coverslips remained unchanged. The cleaned glassware was discarded after 5 days.

Finally, to create the flow channels for each receptor concentration, we used two different methods: parafilm stripes and double-sided polyimide tape. For the initial experiments, after blow drying the glassware with nitrogen, a hand-cut parafilm template was used to create 7 rectangular flow channels of (1 × 22) mm. This was put between the cover slip and the glass slide, and later molten at 125C to attach the glassware together. For the second method, we replaced the molten parafilm with templates of double-sided polyimide tape (50mm width) from Micro to Nano. With the flow channels established, we injected through SUVs, and incubated the sample at non-controlled room temperature for 45 mins. After removing any SUV excess using TAE NaCl buffer, the DNA receptors (Rec) were added in fixed concentrations to each flow channel. To do so, we used Linne (2022) characterization for receptor surface density σ_R, establishing the relation between desired DNA receptors per μm^2 as $\sigma_R = \frac{\#\ receptors\ for\ binding}{surface\ area\ [um^2\}}$, and the required dilution from stock receptor concentration [M] to obtain a fixed number of receptors over the flow channel area $\mathrm{A}{\sim}22\mathrm{x}10^4\ \mu m^2$. For further information, see Table 3.

To assess the binding probability on the sample, we diluted the stock concentration of DNA nanostars $\rho_s = 10^{-4}\ M$ into Tris Acetate-EDTA-NaCl buffer forming a concentration solution of $\rho_n = 10^{-8} M$, and added it on each flow channel. After 20 mins, we expected the sample to be in equilibrium, then the intensity of bound nanostars was measured for seven different σ_R values, ideally covering the entire adsorption range from the minimum bound density I_{bg} to the maximum bound density I_{max} (Table 3). The first channel Ch 1 is always $\sigma_R = 0\ um^{-2}$, corresponding to background intensity, and the last channel Ch 7 holds a maximum number of receptors per unit area for each experiment ($\sigma_R = 300000 - 3750000\ \mu m^{-2}$).

Desired receptor DNA/um^2	Volume receptor [uL]	Stock concentration receptor [M]	Volume TAE - MgCL buffer [uL]
1000	3.19	1.00E-07	19.9
3000	9.56	1.00E-07	13.5
10000	3.19	1.00E-06	19.9

20000	6.38	1.00E-06	16.7
50000	3.19	5.00E-06	19.9
100000	6.38	5.00E-06	16.7
150000	9.56	5.00E-06	13.5
200000	6.38	1.00E-05	16.7
250000	7.97	1.00E-05	15.1
300000	9.56	1.00E-05	13.5
600000	18.4	1.00E-05	4.6
750000	23	1.00E-05	0
1140000	7	5.00E-05	16
3750000	23	5.00E-05	0

Table 3 Desired DNA receptor density (σ_R) receptors/µm^2 employed per channel in our experiments. The values are calculated based on the characterization conducted by Linne (2022). The table shows the relationship between the desired DNA receptor density per μm^2 and the required dilution from the stock receptor concentration [M] using TAE - MgCl buffer, to achieve a fixed number of receptors over the flow channel area (A ~ 22 x 10^4 μm^2).

2.2 Data acquisition using TIRF Microscopy

With the sample ready, we proceed to visualize the adsorption of nanostars onto the SLB, and assess the bound fraction per receptor density, using Total Internal Reflection Microscopy (TIRF). To do so, we utilized an inverted fluorescence microscope (Nikon Ti2-E). The images were taking by using azimuthal TIRF/FRAP illumination module (GATA systems, iLAS 2) and a 100× oil immersion objective (Nikon Apo TIRF, 1.49NA). Each DNA nanostar was labelled with an Atto488 dye, while each receptor had a Cy3 dye. We used the laser excitation wavelength of 561 nm to allocate only the receptors attached to the surface on the sample (TIRF561), then later used the 488 nm wavelength to only excite the nanostars bound to the receptors and take the images there (TIRF488). The emitted fluorescent signal was detected using an EM-CCD Andor iXON Ultra 897.

After identifying the receptors on the surface using TIRF561 starting on Ch 7, the image acquisition protocol was modified for capturing 10 images at 5 locations employing Multi-Dimensional Acquisition (MDA) functionality in MetaMorph software. The image parameters used were TIRF488 laser power 15%, exposure time 50ms, gain 50 and ellipse 100nm. In this approach, we captured five consecutive images at each of the five locations by automatically progressing to the next position after each picture. This cycle was repeated until obtaining a total of 50 images for each channel.

2.3 Data Analysis

2.3.1 Intensity values into bound fraction

We computed the average intensity value from each 50 images per flow channel. The obtained data were then normalized per experiment, utilizing background I_{bg} and maximum intensity I_{max} values (Equation 2.1), to compute the bound fraction per receptor density. The normalization process was conducted in Excel and ImageJ, employing a consistent Region of Interest (ROI) selected across all images. The images were uniformly sized at 512x512 pixels and encoded in 16-bit.

The normalized intensity values were processed in Python, based on the analysis code used by Linne (2022). Changes were introduced in the code to adapt it for the purpose of our current research. The initial aim was to calculate the relation between bound fraction and intensity established by Linne (2022):

$$\Theta = \frac{I - I_{bg}}{I_{\max} - I_{bg}} \tag{2.1}$$

Where the calculation of experimental error involves determining the standard error of the mean (SEM), considering both the number of measurements and their standard deviation. The reported error is based on replicates conducted under the same experimental conditions for each design. By having the bound fraction Θ for each experiment, we then extracted a fitted function Θ_{fit} that represents it, to calculate and plot the superselectivity parameters α and $< n >$ with respect to the receptor density σ_R following Frenkel's superselectivity equation (Eq. 1.1 - 1.5).

2.3.2 Image analysis for clustering

In order to analyze the cluster formations observed with Design C, we developed a Python code to process the TIRF images obtained during the experimental setup, as shown in Figure 8. Initially, we utilized a macro in ImageJ to create masks of the TIRF images, highlighting the locations of the clusters. These masks were then used to extract the size and intensity values of each cluster, which were stored in a CSV file. It's worth noting that the calibration length scale of our KNIC TIRF microscope, which is set to 8 pixels per micrometer, was applied during this process. Subsequently, the generated CSV file serves as a parameter for the Python code. The code then analyzes the CSV file and calculates parameters such as size, count, and intensity values for the clusters per channel. Finally, the code generates plots based on these calculated values. For further information, refer to Section Results.

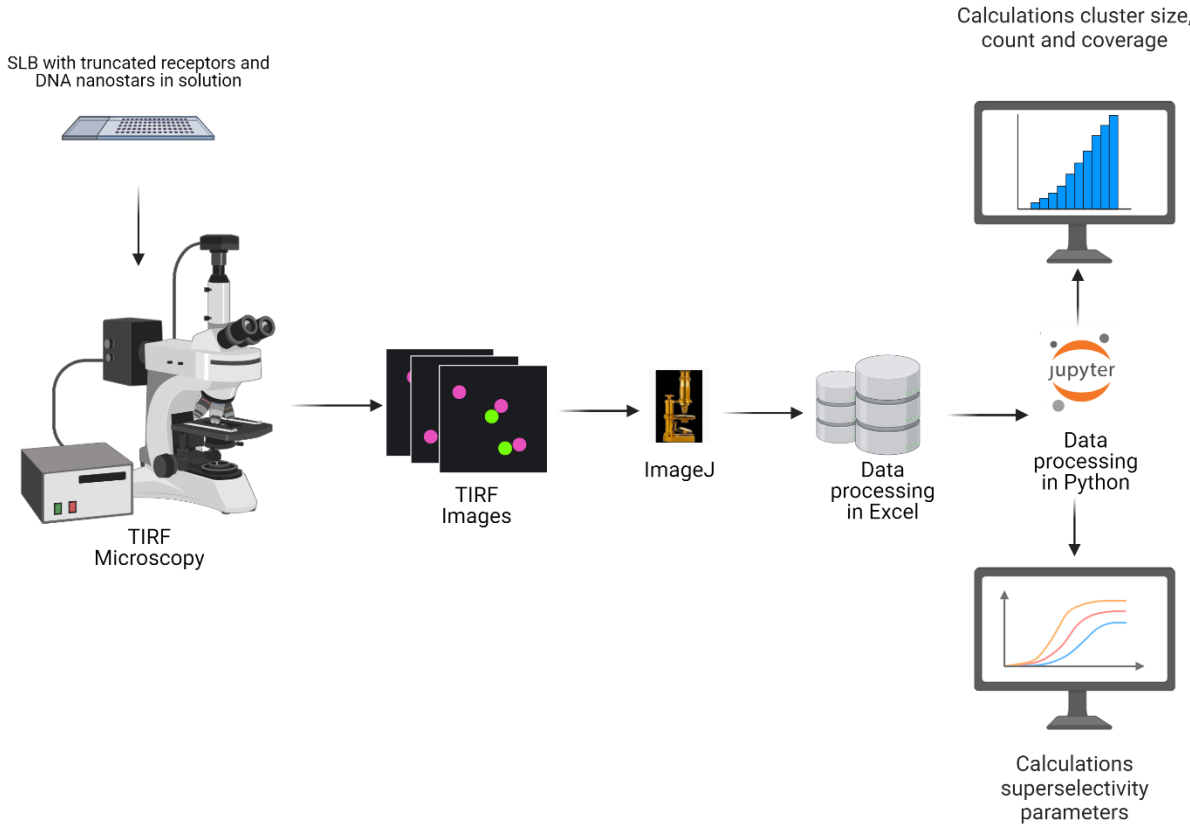


Figure 8 Experimental Workflow Followed in Thesis. This figure illustrates the experimental workflow used in this thesis for analyzing superselectivity using DNA nanostars and receptors. Initially, the sample is placed on a functionalized glass slide holding the DNA nanostars and receptors. The slide is then inserted into a TIRF microscope, where 50 pictures are captured per channel, across seven channels on each glass

slide. The captured images are verified in ImageJ to ensure they are well-focused. From these images, the average intensity value of each picture is extracted, and the raw data is exported into an Excel sheet. In Excel, the data is normalized using the bound intensity equation 2.1. The normalized data is then processed in Python to plot the experimental bound fraction, create a fit bound fraction, and calculate superselectivity parameters. For Design C, there is an alternative data processing step after Excel, where the data is input into a Python code for cluster identification in terms of size, count, and coverage on the sample. Figure created with BioRender.com.

2.3.3 Fitting calculations for binding fraction θ_{fit}

The fit function θ_{fit} adapted from Linne (2022), takes the experimental values for the binding fraction, Θ, and adjusts its values to fit a least square fit denoted as θ_{fit}. The code uses the theoretical definition:

$$\theta_{fit} = \frac{\rho * K_A^{av}}{1 + \rho * K_A^{av}}$$

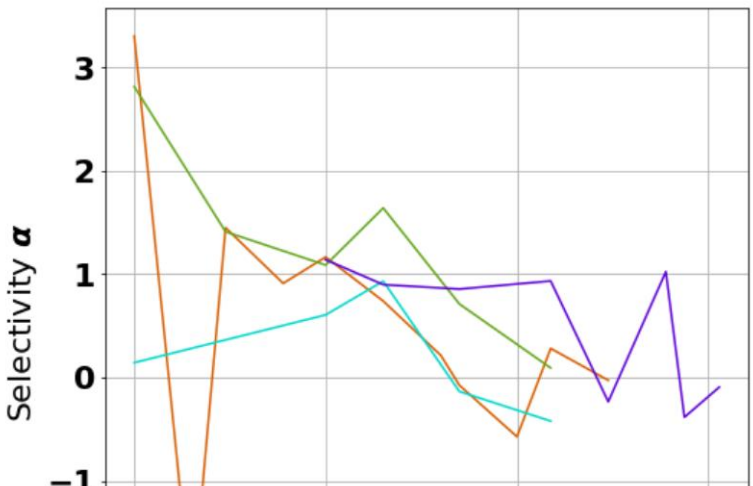
where $\rho_n = 10^{-8}\ M$ is the concentration of DNA nanostars in solution and remains constant for all calculations. The average affinity constant K_A^{av} is calculated as:

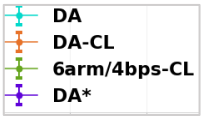
$$K_A^{av} = C\left(\left(1 + \frac{\sigma_R A}{D}\right)^k - 1\right)$$

In this equation A = 24000000 μm^2 represents the total surface area of a glass slide, and k = 6 is the valency of the DNA nanostar. This function optimizes the parameters C and D to minimize the difference between the experimental data θ and the fitted function θ_{fit}. Once calculated, C and D are constant for each calculated curve.

By providing estimates for these parameters, K_A and K_{intra} can be evaluated for each experimental condition given $C = \frac{K_A}{K_{intra}}$, $D = \frac{1}{K_{intra}}$. This approach enables the comparison of the roles of K_A and K_{Intra} in the described superselective behavior across different experimental designs. The obtained θ_{fit} function is then used to calculate superselectivity parameter α, and average number of bound arms <n>, as described in the theoretical model in Section 1.3.

In the context of our study, employing a fitting approach to calculate parameters such as α is essential for several reasons. This approach offers several advantages over direct calculation methods. First and foremost, fitting the experimental data to a mathematical model enables the extraction of α, K_A and K_{intra} parameters with reduced noise and increased accuracy. From our research, the fitting approach mitigates the influence of random fluctuations or measurement errors that can lead to large differences in calculated parameters. Our decision to adopt a fitting method for calculating alpha values was further validated through comparative analyses with direct calculation methods, employed for DA-CL (orange), 6amr/4bps (green) and DA (cyan) datasets, as shown in Figure 9. These analyses revealed that reproducing α values without a fitting approach resulted in significantly higher levels of noise and led to distinct maximum α values compared to the fitting approach, as we will see in the Results section. In this sense, past studies conducted by researchers at TU Eindhoven also utilized a fitting method to determine alpha values from experimental data in weak multivalent interactions (Scheepers et al., 2020), supporting our decision to use a fit.





Receptor density / μm^{-2}

Figure 9 Calculation of Parameter α Without Fitting θ. α calculations without fitting θ for experimental datasets and Linne (CL) data. Cyan and purple lines represent experimental data for Design A (DA) and Design A (DA*), respectively. Orange and green lines correspond to CL data for DA-CL and 6arm/4bps-CL datasets, respectively. The absence of θ fitting results in increased noise levels when compared with fitted functions, emphasizing the importance of fitting for accurate parameter determination.*

2.3.4 Mann Whitney U test for experimental comparison

The Mann-Whitney U test is a non-parametric statistical test that can be used to determine if there is a significant difference between two independent, non-normally distributed groups of data, with a small sample size (less than 15 observations per group) (Frost, 2023). The test ranks all the observations from both occasions and then calculates the sum of ranks for both. If the sums are significantly different, it suggests that the distributions of the two groups are different. The test provides a p-value, which indicates the probability of observing the data if the null hypothesis (that there is no difference between the groups) is true. If the p-value is below a predetermined significance level (~ 0.05), the null hypothesis is rejected, indicating that there is a significant difference between the groups. It is important to note, however, that a significant result does not necessarily imply that the underlying phenomena represented by the two sets of measurements are different. Variations in experimental conditions or other factors may also contribute to observed differences. Therefore, while the Mann-Whitney U test serves as a valuable analytical tool, it is best employed as part of a comprehensive analysis strategy, as we will discuss further in the subsequent chapter.

3 Results

3.1 Experimental Validation using Design A

To begin the study of superselectivity employing 6-arm DNA nanostars, the experimental protocol developed by the Laan Lab was initially verified. This validation utilized the design previously employed by Linne (2022), consisting of 6-arm/6-bp sticky end DNA nanostars with complementary receptors of 6-bps, denoted as Design A (DA) in this chapter (see Figure 10). The hypothesis of this thesis suggested that replicating the existing protocol with identical samples would yield consistent experimental outcomes for superselectivity compared to Linne's findings (denoted as DA-CL). This validation aimed to quantify the variability in these superselectivity results and extend them to new DNA nanostar designs. Additionally, the calculated binding constants K_A, K_{intra}, and α from the obtained results served as confirmation methods for the validation.

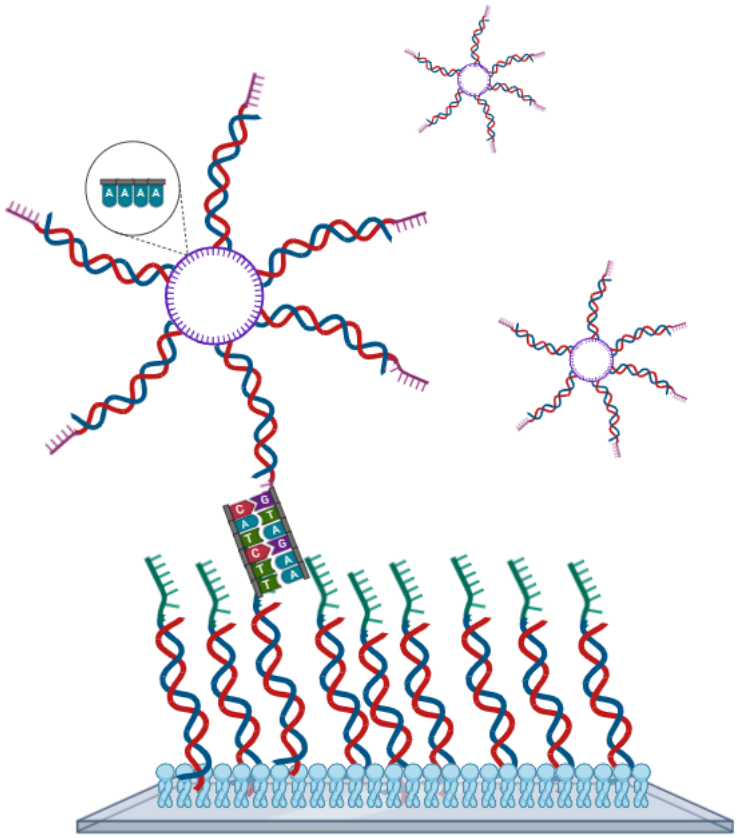


Figure 10 Schematic representation of experimental conditions for Design A. The star-shaped figures represent DNA nanostars (top), interacting with DNA receptors on the target surface (bottom). The receptors are anchored to a supported lipid bilayer (SBL) depicted in blue, which is placed on a microscope glass slide. Each vertex of the DNA nanostar contains a polynucleotide sequence of four adenines between each arm, highlighted by the zoomed-in section. The brush-like structures on both DNA nanostars and receptors symbolize complementary sticky ends, as indicated by the specified interacting nucleotide sequences. This schematic is not to scale. Figures created with BioRender.com.

The experiments presented in this chapter were performed to gain an understanding of the practical details of the protocols involved in DNA nanostar and receptor hybridization, glass slide cleaning, reactivation of Small Unilamellar Vesicles (SUVs) of DOPC, DNA functionalization of supported lipid bilayers, and data acquisition using TIRF microscopy (for detailed procedures, refer to Materials and Methods). Specifically, the aim was to determine the consistency of superselective behavior across different experimental conditions. These changes in experimental conditions entail minor modifications in the necessary protocol adjustments and the learning process. The protocol

modifications included minor alterations in glassware cleaning and the method and materials used for flow channel functionalization with lipid membrane and receptors (see Materials and Methods). The protocols for DOPC SUVs, DNA nanostar and receptor hybridization remained unchanged from the original work by Linne (2022).

Our goal for this experiment was to determine if 6-arm/6-bp DNA nanostars (DA) exhibit comparable adsorption numbers for various densities of chosen DNA receptors (σ_R) compared to the reported findings (DA-CL). To achieve this, we conducted two experimental replicates. To visualize the fluorophores attached to the adsorbed nanostars on the SUVs on the glass slides, and to estimate the number of nanostars bound to the surface receptors, we performed a fluorescence-based assay using TIRF microscopy (Section 2.2). We assumed that changes in fluorescence intensity measured depending on receptor density could indicate alterations in the number of nanostars bound, with higher fluorescence intensity values expected to correlate with increased binding of nanostars to the surface, as shown in Figure 11.

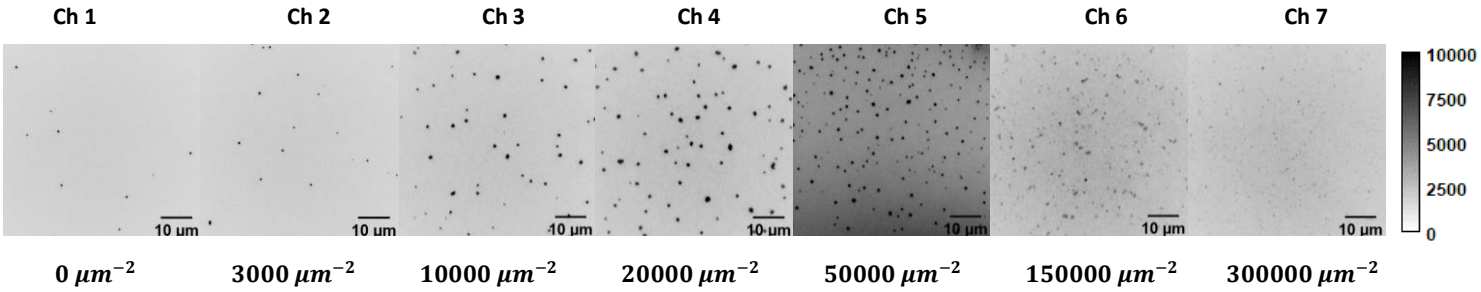


Figure 11 TIRF Microscopy Images for seven channels on the Target Surface for Design A (Exp #1). Images captured via TIRF microscopy representing seven channels on the target surface for Design, experiment #1. From left to right: Channel 1 serves as the negative control ($\sigma_R = 0 \mu m^2$) establishing the background signal I_{bg}. Channel 7 exhibits the maximum receptor density and ideally should display the highest registered intensity I_{max}, but it was consistently not observed across experiments. The images correspond to the following receptor densities on each of the seven channels: σ_R = [0, 1000, 10000, 20000, 50000, 150000, 300000] μm^2. All images were adjusted to have the same intensity values for comparison, from 0 to 10000 intensity units.

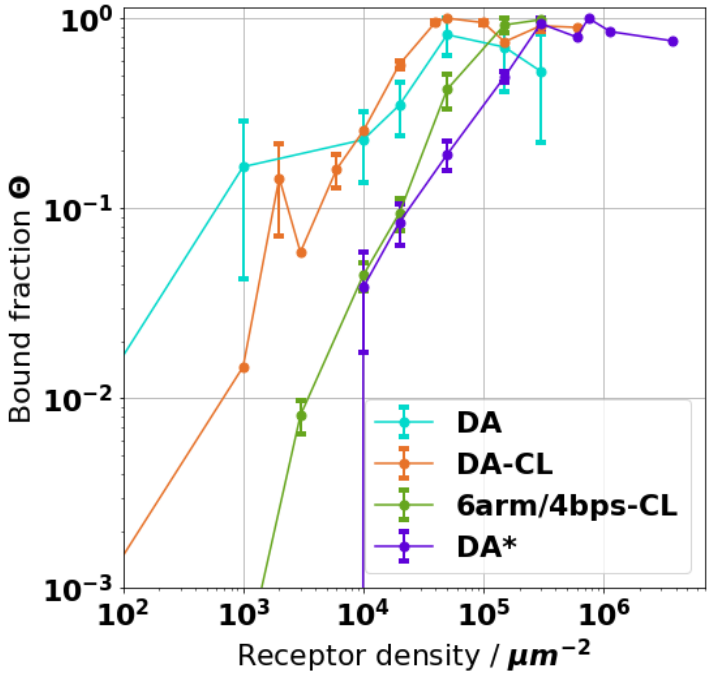


Figure 12 Experimental measurements of bound fraction θ as a function of receptor density for Design A and Design A (with a mismatch in the sticky end sequence between DNA nanostars and receptors, further explained in section see section 3.2), compared to experimental results*

obtained by Linne (2022) (DA-CL, 6arm/4bps-CL). Each curve represents the results of at least 2 experimental replicates per condition. Error bars indicate the standard error of the mean value per receptor density.

To build the sigmoidal bound fraction θ, we sampled seven receptor densities going from a maximum receptor density value (σ_R max) to zero saturation state ($\sigma_R = 0\ \mu m^{-2}$), to cover the full spectrum of adsorption. On each receptor density, we measured the mean intensity value along a fixed region of interest (ROI), standard for all the experiments unless there were image artifacts that could notably interfere with the average value. For the case of DA, the maximum receptor value (σ_R= 300000 μm^{-2}) did not correspond to the highest intensity value registered, in fact it corresponded to σ_R= 50000 $\mu m^{-2}$and σ_R= 150000 μm^{-2} on each of the experiments, as seen in Figure 12. The receptor density where the sample reached a fully adsorbed state corresponds to the highest normalized intensity registered (Θ ~ 1) as seen in Figure 12.

Upon qualitative assessment of the results presented in Figure 12, it is observed that, overall, the distribution of bound fraction obtained falls within the range of error bars, reaching full adsorption around the same range of receptor densities (σ_R = 50000 μm^{-2}), and displaying a reduction in bound fraction after reaching maximum adsorption state, a phenomenon observed across other designs and previous reported values (Linne, 2022) (Figure 6). Our experiments also reported a high bound fraction towards lower receptor densities (σ_R= 1000 μm^{-2}), consistent with previous reports. While acknowledging that visual verification alone may not be sufficient to ensure experiment reproducibility and protocol reliability, we posed the question of whether these experimental curves could be quantitatively assessed for statistical similarity.

3.1.1 Are DA-CL and DA statistically significantly similar?

In comparing the experimental datasets of DA-CL and DA, we utilized the Mann-Whitney U test, a non-parametric method detailed in Methods. This method, advantageous for its flexibility with dataset distributions and receptor density variations, requires formulating a null hypothesis. Our null hypothesis posited no statistically significant difference between the distributions of the DA-CL and DA datasets. A Mann-Whitney U p-value below 0.05 indicates significance, encouraging rejection of the null hypothesis and suggesting a notable difference between the datasets. Conversely, a p-value above 0.05 fails to reject the null hypothesis, indicating similarity between the datasets, implying they represent the same phenomenon.

Calculations with the CL-DA dataset as reference yielded a Mann-Whitney U test p-value for DA of 0.552, suggesting no statistically significant difference between the distributions of DA-CL and DA. This implies that the two datasets may represent the same phenomenon, supporting the reliability of the experimental protocol used for both datasets.

3.1.2 Sensitivity analysis: Calculation of Binding Constants K_A, K_{intra}, and superselectivity parameters θ, α, and <n>

Moving beyond the experimental results for bound fraction, we now evaluate the reliability of the experimental protocol while calculating the binding constants and superselective parameters, by employing a sensitivity analysis. Let us recall that the terms describing superselective behavior, K_a , K_{intra}, α, and <n>, are calculated from fitting the adsorption profile θ using a least square method for nonlinear curves (See Methods), which makes the fit more sensitive compared to fitting a linear curve. For this reason, the sensitivity analysis focused on the assessment of how variations or uncertainties in the input parameters propagate through the calculations to affect the outputs (K_A, K_{intra}, α and <n>).

By systematically varying the input data points within three proposed cases, and observing the resulting changes in the output, the sensitivity analysis provided insights into the relative importance of different sources of uncertainty. Our hypothesis was that the least square fit method used to quantify superselective behavior from experimental data is heavily dependent on the input data points, and that the superselectivity parameters calculated from it are sensitive to the choice of data points for the fit. This turns to be important considering that Linne (2022)

compares DNA nanostars with different valency number (number of arms) based on these calculations, and that this procedure is the decisive method to say if a sample is behaving superselectively or not.

To quantify the variation, three cases were postulated to vary the data points that represent the sigmoidal experimental adsorption profile θ, into the generated fitted version of it, called θ_{fit}, depicted in Figure 13. Case 1 includes all the data points. Case 2 excludes the outliers from the sigmoidal behavior, in this case the low receptor density values, and Case 3 excludes the low receptor density outliers and the data points after full adsorption was reached.

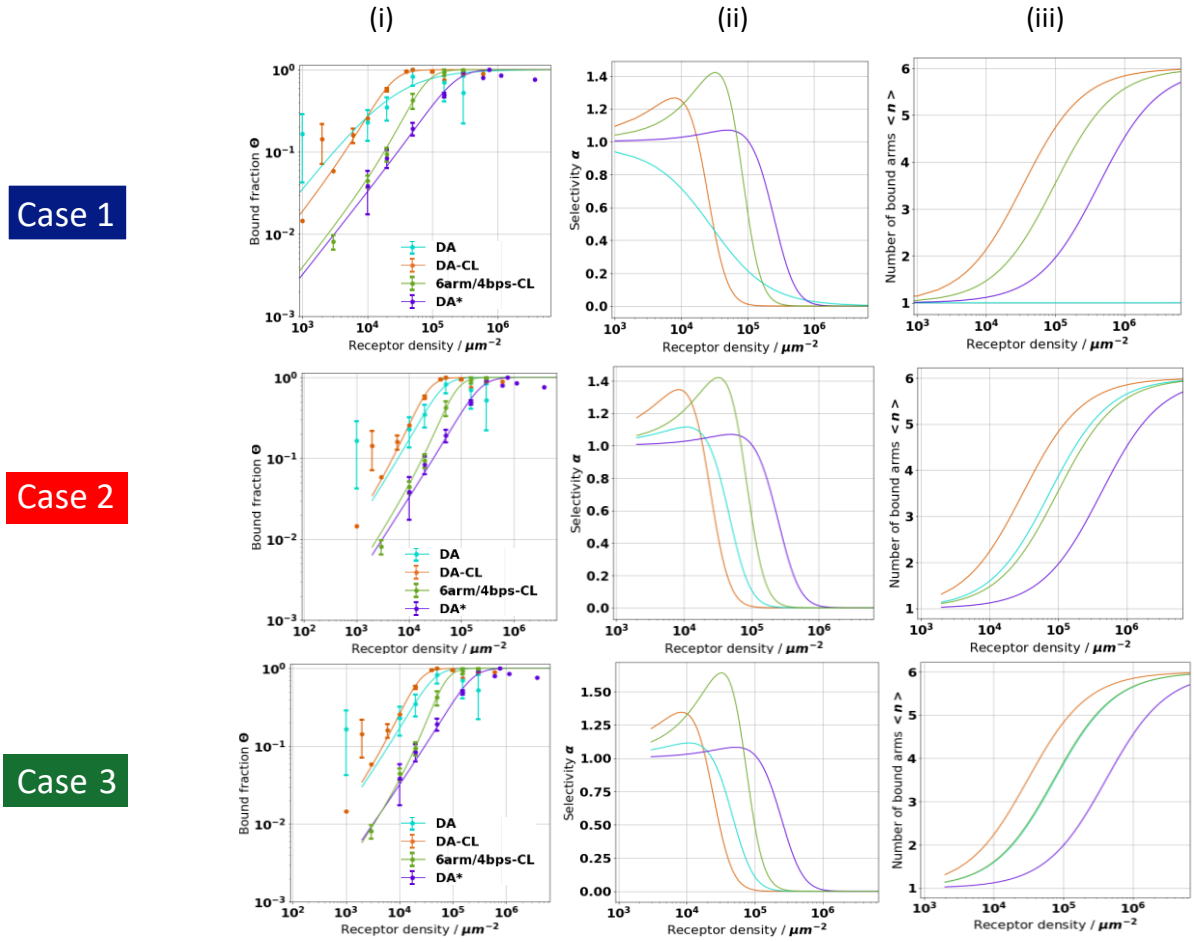


Figure 13 Sensitivity analysis for the calculation of superselectivity parameters across three cases: Case 1 (blue), Case 2 (red), and Case 3 (green). From left to right: (i) Adsorption profile θ_{fit} with respect to receptor density σ_R, (ii) Superselectivity parameter α vs. σ_R, and (iii) average number of bound arms <n> with respect to σ_R. The figure illustrates the sensitivity of the fitting process for superselectivity parameters across three cases. Case 1 considers all experimental data points, resulting in notable differences in the slope of the fitted adsorption profile compared to Case 2 (removing low receptor density outliers) and Case 3 (removing outliers and data points after full absorption). These differences influence the calculated superselectivity parameter α and <n>. Notably, Case 1 shows a deviation from the superselective regime observed in Case 2 and Case 3, resembling monovalent particle behavior. Comparisons with reported values by Linne (2022) reveal discrepancies, particularly in the maximum α value (sometimes referred as α_{max}) and <n>. Case 2 and Case 3 exhibit similar trends in α value and <n>, indicating the robustness of the fitting process under outlier removal. These findings highlight the importance of considering variations in the fitting approach when analysing superselectivity parameters.

From Figure 13 the most important differences observed along the three proposed cases for DA and DA-CL are (1) differences in the slope for the fitted adsorption profile θ_{fit} , (2) differences in the calculated max value peak for superselectivity parameter α, and (3) differences in the calculated average number of bound arms <n>.

In Case 1, the slope value from DA was mostly affected by the inclusion of all the experimental data points, leading to a small reduction in slope values compare to DA-CL. From this, it is evident that the model is highly sensitive to slight slope changes as the calculated α value for this design describes as it the adsorption profile was no longer superselective, resembling the results describing the adsorption for monovalent particles with low free binding energies (Martinez-Veracoechea & Frenkel, 2011). Moreover, this also affects the calculation of the average number of bound arms, confirming the interpretation of the monovalent particle, as only one arm being bound. If we compare to the values reported by Linne (2022) (see Figure 13) for DA-CL, the max value for the superselective parameter α reported in Case 1 differs slightly, with DA-CL $\alpha \sim 1.27$ and DA $\alpha \sim 0.94$ (Table 4). However, the previous reported value for <n> seems to differ, as for this $\sigma_R \sim 10^5$, <n> ~ 5 arms, compared to the reported <n>~4 arms for the same receptor density value.

In Case 2, the removal of the low receptor density outliers allowed both DA and DA-CL slopes to behave more similarly, although the calculated α values for both vary significantly. Unlike Case 1, the shape of the α function allow to conclude that both describe multivalent superselective adsorption (Martinez-Veracoechea & Frenkel, 2011). Comparing with the samples reported by Linne (2022) (see Figure 2), DA-CL for this case registered a slightly higher α max value (1.35), and DA values although in the superselective regime (1.11), are lower than DA-CL. Notably, for DA-CL $\sigma_R \sim 10^5$, the reported <n>~4 arms, as for our calculations this value corresponds to <n>~ 5 arms. For DA, <n>~4 arms, which then coincides with the previous studies.

Lastly, for Case 3, the removal of both low receptor density outliers and experimental data after reaching full absorption shows little to no change in the slope values for DA-CL and DA seen in Case 2. Superselective values do not change compared to Case 2. In consequence, the curves for <n> did not change either.

The results obtained showed Case 1 to be the closest to the reported values by Linne (2022), but also the case where the difference in the max value of α is the largest between DA-CL and DA. To continue the exploration of superselective parameters and how the fit used to calculate those could affect the results, a second design was explored. For this case, a change in the free binding energy available for ligand-receptor interactions was introduced.

3.2 Experimental Effects of Complementary Sequence Variations in sticky end using Design A*

To continue our study of superselectivity using 6-arm DNA nanostars, we created a new design where the complementary receptor's sticky end does not fully coincide with the sticky end of the ligands. The design consists of a 6-arm/4-bps sticky DNA nanostar with a 6-bps sticky end receptor on the adsorption surface, as depicted in Figure 14. This new design will be referred to as Design A* (DA*).

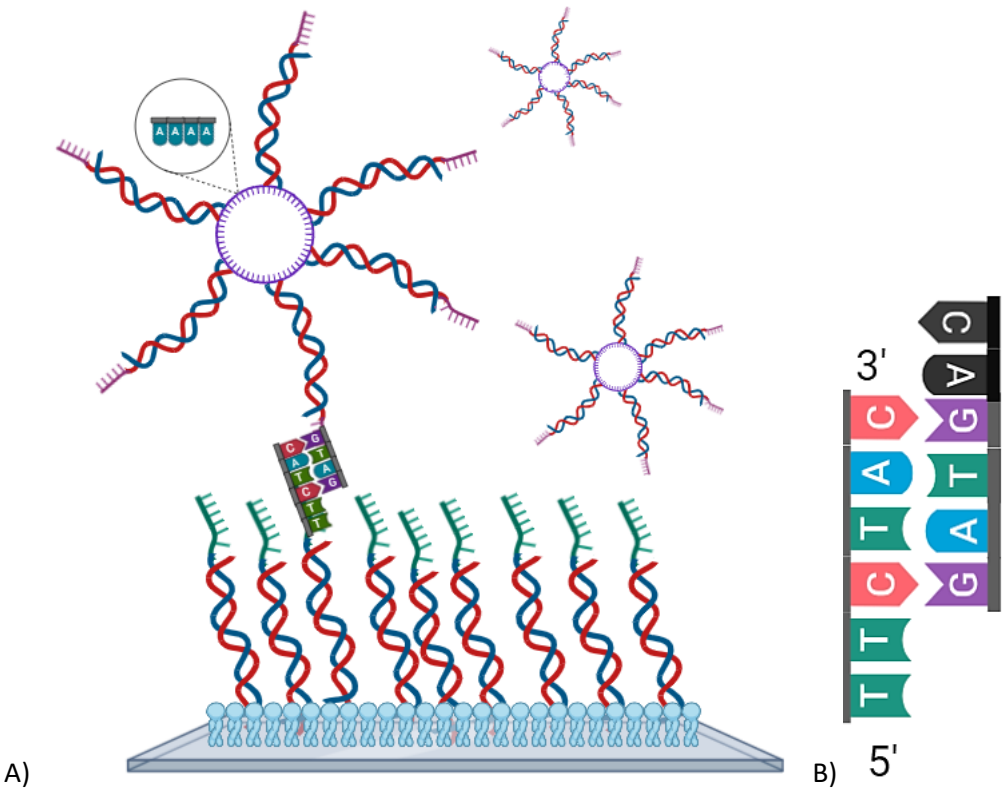


Figure 14 Schematic of Design A. A) Schematic representation of the Design A* DNA nanostar with its associated DNA receptor. The nanostar is shown interacting with a DNA receptor on the target surface. The polynucleotide sequence of four adenines between each arm of the DNA nanostar is detailed. Brush-like structures on both the DNA nanostar and the receptors symbolize the complementary sticky end mismatch, which are specified with nucleotide sequences. B) Close-up of the sticky end nucleotide mismatch interaction: the left sequence (5' to 3') corresponds to the sticky end on the receptor strand, while the right sequence corresponds to the sticky end on one of the DNA nanostar arms. The nucleotides marked in black belong to the DNA nanostar. Figures created with BioRender.com.*

We hypothesized that such "mismatch" in the complementary sequences for the sticky ends would increase superselective parameters as a decrease in the affinity between ligands and receptors corresponds to one of the reported rules for increasing superselective behavior (Dubacheva et al., 2023). Thus, the objective of these experiments was to evaluate the effects of a change in free binding energy, translated into the sequence mismatch, into the superselective behavior of our proposed probes. In the same way, we wanted to evaluate quantitatively the effect on the binding strength constants K_A, K_{intra}, α and <n>, and compare it to our results DA, and previously reported results DA-CL and 6arm/4bps -CL from Linne (2022).

In terms of the protocols used, the same changes used for the glassware cleaning preparation for Design A were used, and the method and materials used for the flow channel functionalization with lipid membrane and receptors (see Methods). The protocols for DOPC small unilamellar vesicles (SUVs), and DNA nanostar and receptor hybridization were kept like the original works.

We proceeded to estimate the DNA nanostars adsorbed to the surface with fluorescence-based assay and used TIRF microscopy to visualize the fluorophores attached to the adsorbed nanostars on the glass slide (see Figure 15), and reconstruct the experimental adsorption profile θ with respect to receptors (σ_R), as shown in Figure 12. The same assumptions about the direct correlation between intensity and number of nanostars bound were employed.

For each experiment, we imaged seven receptor densities and measured the mean intensity value along a fixed region of interest (ROI) of 4096 μm^2. To build the sigmoidal adsorption profile for this design, we were required to sample much higher receptor densities than the employed for Design A to cover the full spectrum of adsorption. Moreover, again the maximum receptor value (σ_R = 3750000 μm^2) did not correspond to the highest intensity value registered, in fact it corresponded to σ_R = 525000 μm^2 and σ_R = 750000 μm^2, as seen in Figure 12. Using this design, a decrease in bound fraction after the full absorption was sampled, a phenomenon observed for DA and previous reported experiments (Linne, 2022) (Figure 5).

From a graphical assessment of the adsorption profile, we note two of the low receptor density values for bound fraction for DA* coincide with values reported for the 6arm/4bps-CL curve. Nonetheless, for increasing receptor densities, there is a clear reduction in bound fraction values for DA*until both curves reach saturation at σ_R= 300000 um^{-2}. Since 6arm/4bps-CL was not sampled for receptor density values over 10^5, we do not know if it could have shown the characteristic reduction in bound fraction as DA*, as seen in Figure 6**Error! Reference source not found.**.

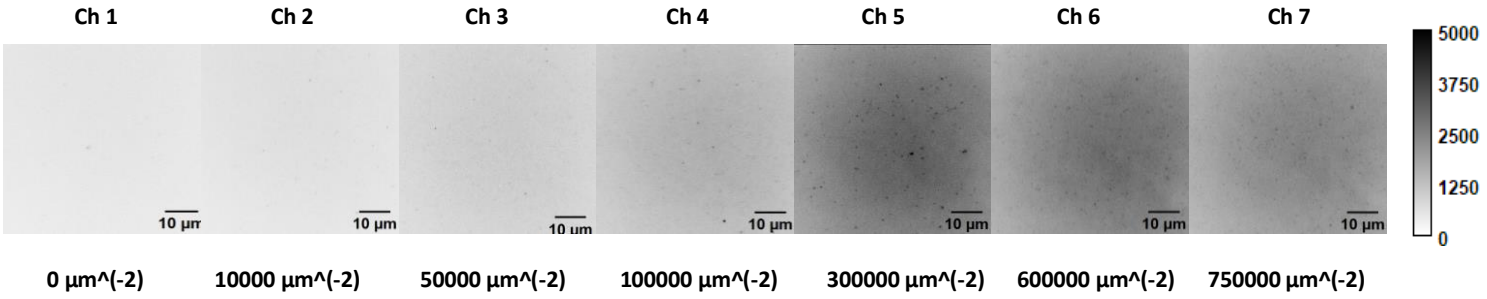


Figure 15 TIRF Microscopy Images of Design A on the Target Surface. From left to right: Channel 1 serves as the negative control ($\sigma_R = 0\mu m^{-2}$) and defines the background signal I_{bg}. Channel 7, intended to have the maximum receptor density and ideally the maximum registered intensity, although I_{max} did not consistently achieve this across experiments. The images were captured for varying receptor densities on each of the seven channels per experiment, using all the possible densities σ_R = [0, 10000, 50000, 100000, 300000, 525000, 600000, 750000, 1140000, 3750000] μm^{-2}. The intensity ranges shown go from 0 to 5000, as shown in the heatmap column on the right.*

Proceeding with the calculation for superselectivity values and binding constants using the Sensitivity Analysis presented in 3.1.2, we compare the values obtained for each of the designs used in Figure 16. Table 4 presents the values in Figure 13, along with the calculated error values. For each table, the colour represents the cases illustrated in Figure 13 and Figure 16.

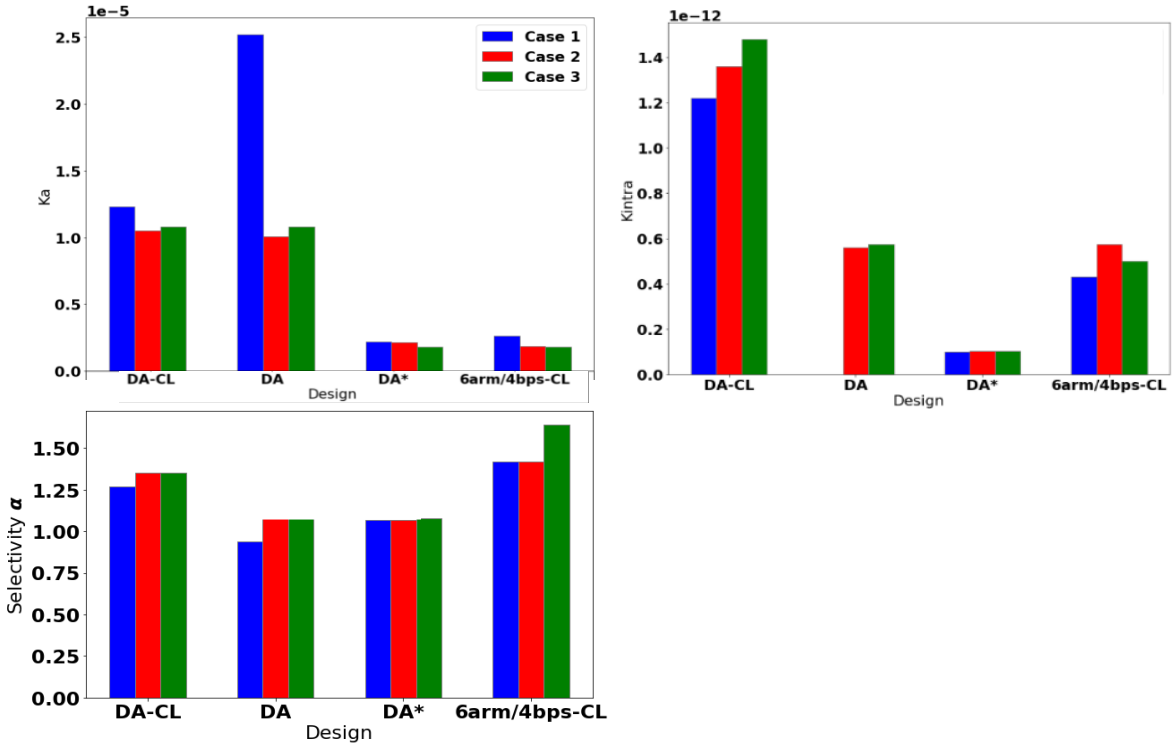


Figure 16 Comparative Boxplot Analysis for Ka, Kintra, and α_max Values. This figure presents boxplot graphs comparing the values of K_A, K_{intra}, and max selectivity parameter α_{max} for Design A, Design 6arm/4bps-CL (DA-CL) sourced from Linne (2022), Design A (DA), and Design A (DA). The analysis includes three sensitivity analysis cases, distinguished by colors: Case 1 (blue), Case 2 (red), and Case 3 (green).*

CASE 1 Design	K_A $(M^{-1} * 10^{-5})$	K_{intra} $(* 10^{-13})$	K_A/K_{intra} $(M^{-1} * 10^{7})$	α_{max}
DA-CL	1.23 (1.48)	12.2(6.26)	1.01	1.27
DA	2.52 (1.40)	3.49 10^{-5} (1.37)	7210	0.94
DA*	0.22 (0.48)	0.99(1.0)	2.24	1.07
6arm/4bps-CL	0.27(0.09)	4.31(0.58)	0.62	1.42

CASE 2 Design	K_A $(M^{-1} * 10^{-5})$	K_{intra} $(* 10^{-13})$	K_A/K_{intra} $(M^{-1} * 10^{7})$	α_{max}
DA-CL	1.05(1.52)	13.6(7.96)	0.773	1.35
DA	1.01(5.43)	5.59(13.5)	1.80	1.11
DA*	0.22(0.10)	0.99(1.00)	2.24	1.07
6arm/4bps-CL	0.27(0.09)	4.31(0.06)	0.68	1.42

CASE 3 Design	K_A $(M^{-1} * 10^{-5})$	K_{intra} $(* 10^{-13})$	K_A/K_{intra} $(M^{-1} * 10^{7})$	α_{max}
DA-CL	1.05(1.55)	13.6(8.14)	0.773	1.35
DA	1.01(5.43)	5.59(13.5)	1.80	1.11
DA*	0.22(0.05)	1.02(1.12)	2.10	1.08
6arm/4bps-CL	0.19(0.003)	5.75(0.03)	0.33	1.62

Table 4 Calculated values for K_A, K_{intra}, $\frac{K_A}{K_{intra}}$ and α_{max} for three cases for sensitivity analysis, derived from optimization terms from fit as described in Methods. In parentheses, the calculated propagated errors from the optimization parameters. For more information on error calculation, see Appendix (Section 7.3).

Let us recall that both DA-CL and DA have a sticky end length of 6 base pairs. In contrast, 6arm/4bps-CL and DA* DNA nanostars feature a sticky end length of 4 base pairs. Despite these differences, all four designs share the same valency. We anticipated significant variation in K_A and K_{intra} among the different sticky end lengths but expected less variation within each group.

We expected similar K_A values for DA and DA-CL since they represent the same case. Similarly, comparable values were anticipated for DA* and 6arm/4bps-CL. K_A represents the interaction strength between ligands and receptors for the initial binding event, depending on the number of base pairs on the sticky ends in our model. Larger K_A values shift the critical point of the selective interaction curve θ to lower concentrations, potentially reducing superselective behavior. Notably, there is a one-order-of-magnitude difference between the groups, with DA and DA-CL exhibiting early jumps to superselectivity, reflected in their K_A values. Conversely, DA* and 6arm/4bps-CL consistently show the lowest K_A values, with the transition point to superselectivity occurring at higher receptor densities. Case 1 exhibits the highest variability in K_A values among the designs, while Cases 2 and 3 show the lowest variability. The observation of higher K_A for DA in Case 1 supports the explanation of monovalent binding, where the initial binding event is strong enough to inhibit subsequent multibinding.

For K_{intra}, we expect to see a higher K_{intra} constant for higher superselective values, indicating more ligands are bound to the receptors on the surface after the first bound has been established. We observe DA and DA-CL tend to have the highest K_{intra} values registered, indicating these designs use more arms to bind than the other two designs. It is also noted that the K_{intra} value from DA varies considerably compared to DA-CL in all cases, which ideally should not happen. For Case 1, the Kintra value obtained is surprisingly small, indicating a possible problem with the fit used to represent this function. In all cases, DA* consistently shows the lowest K_{intra} values, possibly indicating that the sticky end mismatch is interfering with the binding process of more than one ligand.

When put together, the $\frac{K_A}{K_{intra}}$ ratio is ultimately what dictates a design is more or less superselective. We expect to see low ratios for high superselectivity. We observe consistently in all cases 6arm/4bps-CL to have the lowest K_{intra} value, followed by DA-CL. And in all cases, DA* seems to have the highest K_{intra} value. The variability among values through the considered cases is minimum, only for the Case 2 there is an exception for DA that we attribute to the quality of the fit. These results are consistent with the maximum superselectivity registered for the designs.

Lastly, the α_{max} maximum selectivity parameter registered consistently shows 6arm/4bps-CL to be the most superselective design, followed by DA-CL, without large changes across cases. DA* was always the least superselective design.

3.3 Influence of DNA Nanostar Flexibility on Superselectivity Measurements using Design C

To investigate the influence of DNA Nanostar flexibility on superselectivity (Design C), we created a new design by increasing the flexibility and length between the 6-arms. This was achieved by altering the polynucleotide sequence connecting the arms from four adenines to ten thymines, as seen in Figure 17. Following the same protocol used for the previous designs, we observed significant differences in the average intensity values obtained. Specifically, the θ curve exhibited anomalous behavior compared to previous experiments, as depicted in Figure 18.

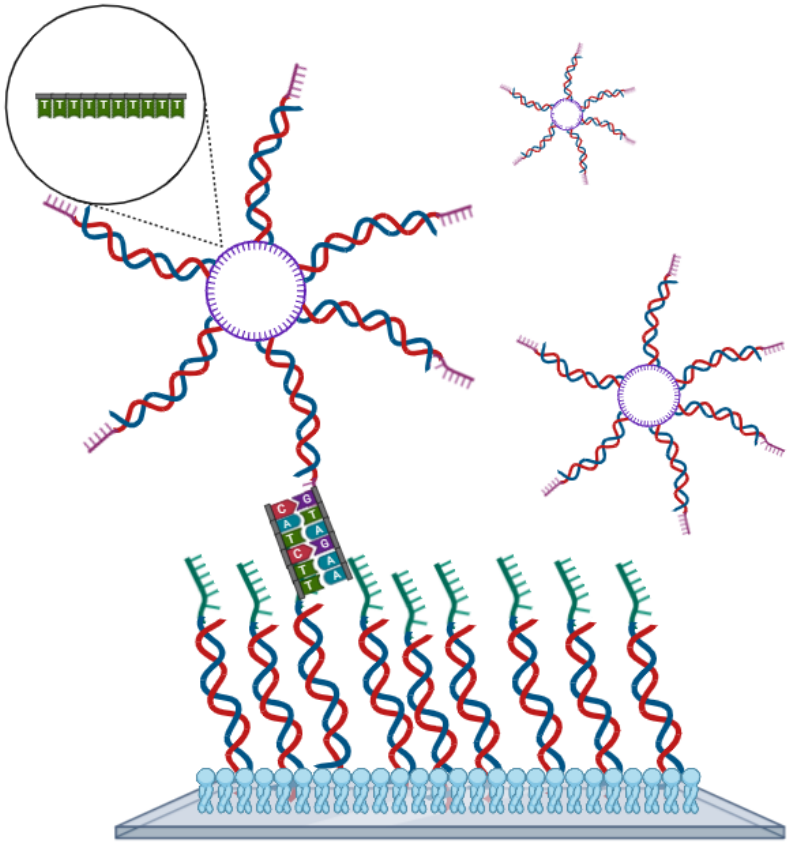


Figure 17 Schematic of Design C of DNA nanostars with their associated DNA receptors. The nanostar is shown interacting with a DNA receptor on the target surface. The polynucleotide sequence of ten thymines between each arm of the DNA nanostar is detailed. Brush-like structures on both the DNA nanostar and the receptors symbolize the complementary sticky ends, which are specified with nucleotide sequences. The scheme is not to scale. Figure made with BioRender.com

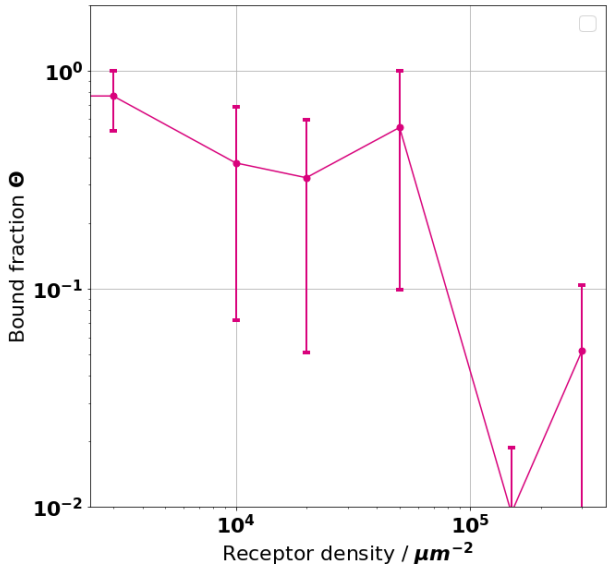


Figure 18 Adsorption profile θ for Design C. Notably, the behavior depicted deviates considerably from previous designs, potentially indicating a decrease in coverage as receptor density increases.

Additionally, we observed the formation of "cluster" structures over some, but not all, of the receptor densities σ_R, meaning larger particles were seen here compared to any of the previous two designs in the TIRF microscopy images. To further investigate and describe this effect and how it could be connected to the increased flexibility in the sample, two different analysis methodologies were created.

Initially, an image analysis algorithm in Python was developed to examine each Total Internal Reflection Fluorescence (TIRF) image individually for the particles formed. This enabled us to analyze the size, intensity, and distribution across each channel within the images. For our experiments, we utilized 7 receptor density levels σ_R, labeled as *flow channels* as follows: Ch1 = 0 μm^{-2}, Ch2 = 3000 μm^{-2}, Ch3 = 10000 μm^{-2}, Ch4 = 20000 μm^{-2}, Ch5 = 50000 μm^{-2}, Ch6 = 150000 μm^{-2}, Ch7=300000 μm^{-2}. Since there was a low correlation between the two replicates done with Design C, the results for each are shown separately.

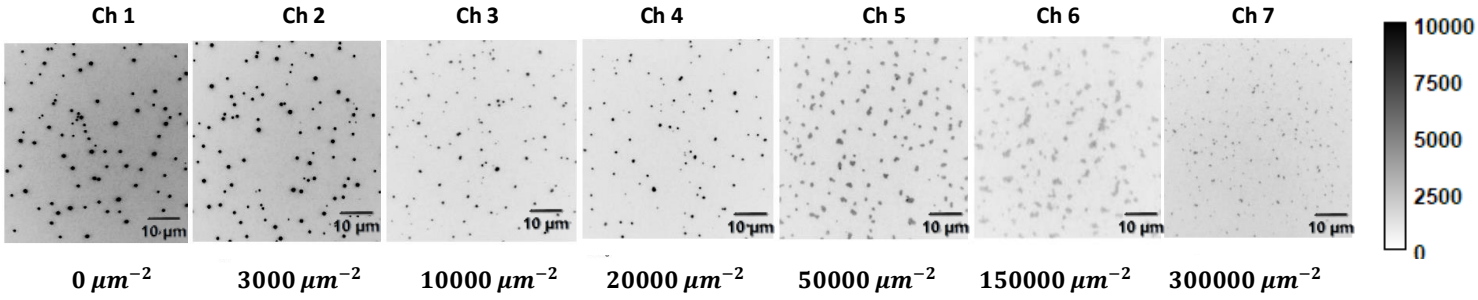


Figure 19 TIRF images of Design C, first replicate. From left to right: Ch1 to Ch7. Cluster formation and intensity values relative to background are observed. Channels 1 and 2 exhibit darker clusters, indicating higher average intensity values compared to other channels. Significant increases in cluster size are noted for Ch 5 and 6, followed by a decrease in Ch 7. Additionally, a decrease in cluster intensity with increasing size is observed. The intensity ranges are scaled from 0 to 10000, as depicted in the heatmap column on the right. Colors are inverted to highlight cluster formations.

For the first replicate, shown in Figure 19, our observations revealed a tendency for the formation of large particles or "clusters" in Ch 5 and Ch 6, which are notably reduced in size in Ch 7, as depicted in the histograms on Figure 20. Our analysis allowed us to visualize the largest clusters formed in Ch 5 and Ch 6, with a 5-fold count increase in Ch 6 compared to Ch 5. Ch 7 shows an important reduction in both size and count with respect to Ch 6. All the other channels show a cluster size no larger than 1-2 um^2 . The observations point at Ch 5 and Ch 6 as the

preferential for these structures since the DNA nanostar concentration in solution was kept constant at $\rho_n = 10^{-8}\ M$ for all the channels.

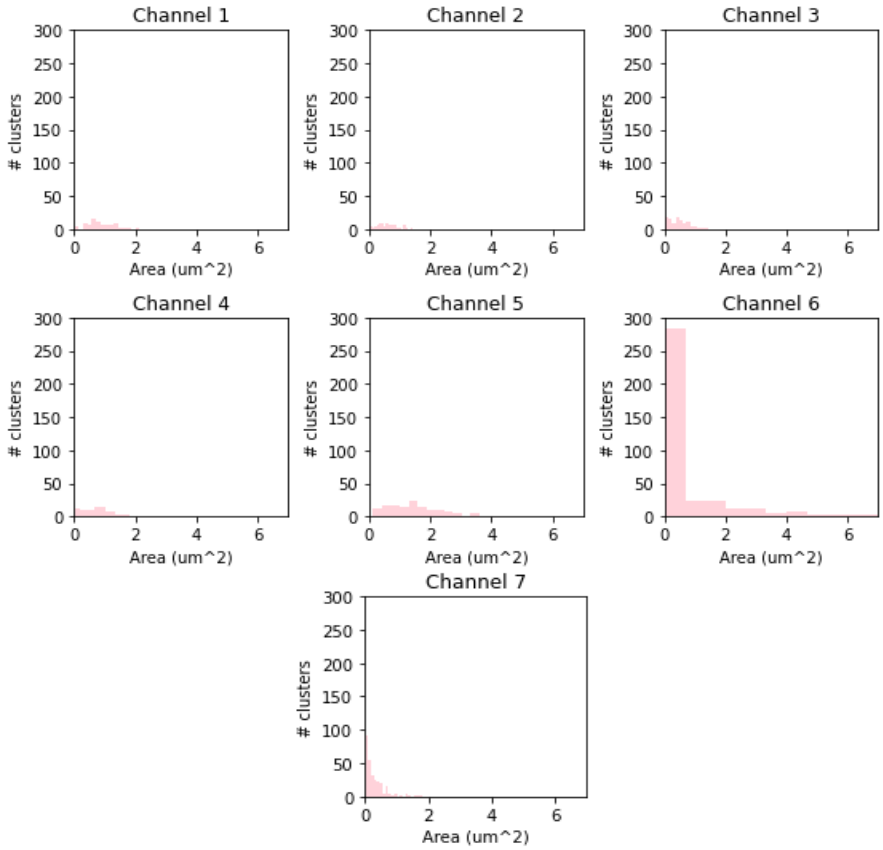


Figure 20 Histograms for cluster counting vs. size for each receptor density concentration, enumerated as flow channels, in replicate 1. Channels 1 to 7 correspond to receptor density concentrations of Ch1 = 0 μm^{-2}, Ch2 = 3000 μm^{-2}, Ch3=10000 μm^{-2}, Ch4=20000 μm^{-2}, Ch5=50000 μm^{-2}, Ch6=150000 0 μm^{-2}, and Ch7=300000 μm^{-2} respectively. Channels 5, 6, and 7 exhibit the largest counts of clusters, with Channel 6, representing the largest clusters and the highest count.

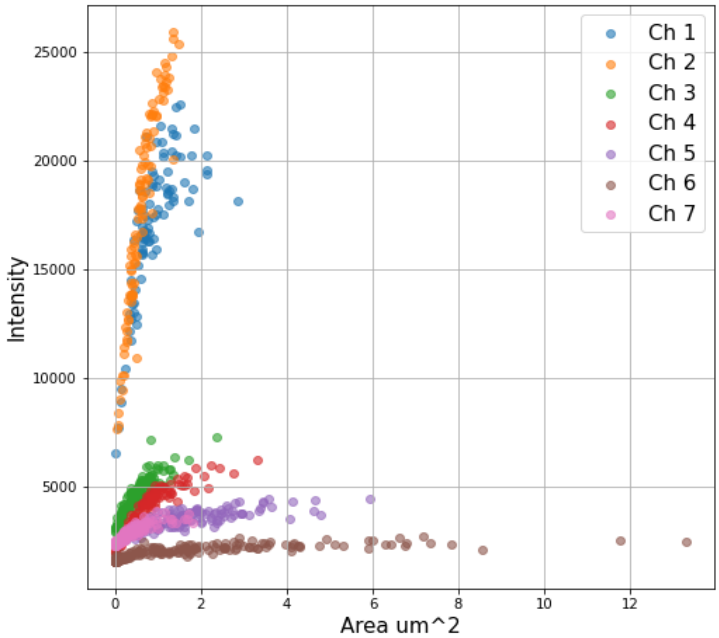


Figure 21 Cluster size and distribution for Design C, first replicate. Scatter plot show the size distribution of the clusters per channel vs the intensity registered during TIRF imaging. Notably, Ch 1 and Ch 2 show smaller, but highly intense clusters, as Ch 6 and Ch 5 show the largest clusters with lowest intensities registered.

Since one of the considerations of the experimental model we are using ignores any cooperation or interactions among ligands (see Section 1.3.1), the formation of clusters can indicate a violation of this assumption. The latter would imply that our readings for intensity values to calculate θ (see Methods) could not be a measurement of coverage of single nanostars adsorbed uniformly to receptors but would represent the adsorption of interacting nanostars or receptors, or both, forming clustered particles. Our hypothesis was that the observed clustering phenomenon could be dominating the read for average intensity values, requiring a reinterpretation for the adsorption profile θ. To do so, our next analysis focused on relating cluster size with intensity value registered. For this replicate, we found a negative correlation between cluster size and intensity value, showing Ch 5 and Ch 6 have the largest clusters with lowest intensity values. In the same way, Ch 1 and Ch 2 show the highest intensity values with the smallest clusters. It is noted these intensity values showed to be around 4.5 times higher than for the remaining channels, nothing alike to any other experiment we have seen before.

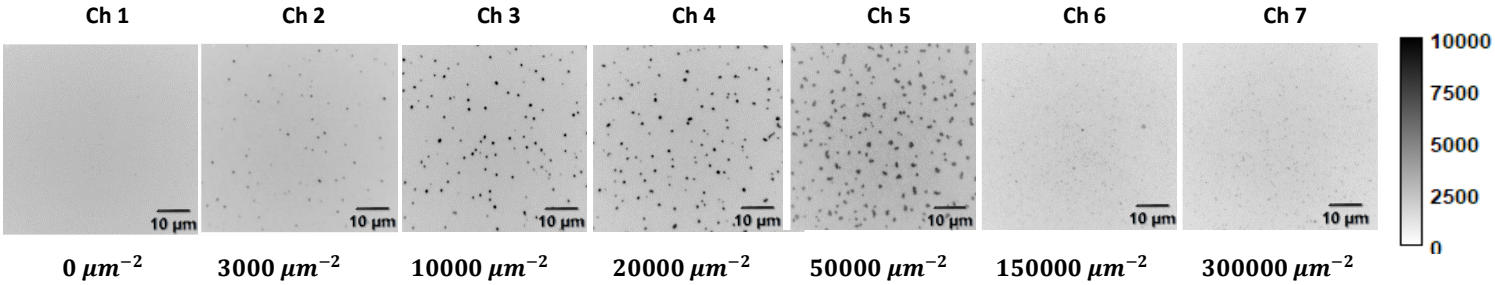


Figure 22 TIRF images Design C, second replicate. From left to right, Ch1 to Ch7. We observe the cluster formation and their intensity value relative to background. We observe Ch 1 and Ch2 show zero to low cluster density. We note the cluster size increases significantly for Ch 4 and Ch 5, to decrease in size and intensity in Ch 6 and Ch 7. Moreover, we observe the cluster intensity to decrease as the size increases. The intensity ranges shown go from 0 to 10000, as shown in the heatmap column on the right.

After repeating the analysis for the second replicate, we observed again the preferential formation of large clusters in Ch 5, with an exceptional higher count for smaller clusters in Ch 6 and Ch 7, as depicted in Figure 23. Overall, the maximum cluster sizes reached here were half the size of the ones registered in the first replicate. Our analysis over size and intensity, as seen in Figure 24, show a negative correlation cluster size vs intensity values for Ch 5. Interestingly, we observe Ch 1 and Ch 2 clusters register low intensity values, different to the first replicate. Nonetheless, we observe Ch 6 and Ch 7 clusters show the lowest intensities, coinciding with the observations for the first replicate. It is important to clarify that intensity values are not absolute across replicates; therefore, comparison is only possible relative to the values found for each experiment. To compare among replicates, and for extension experiments, it is required to normalize the values, as explained in section 2.3. Since we were only interested in the relative increase of intensity, the values were not normalized for these two cases.

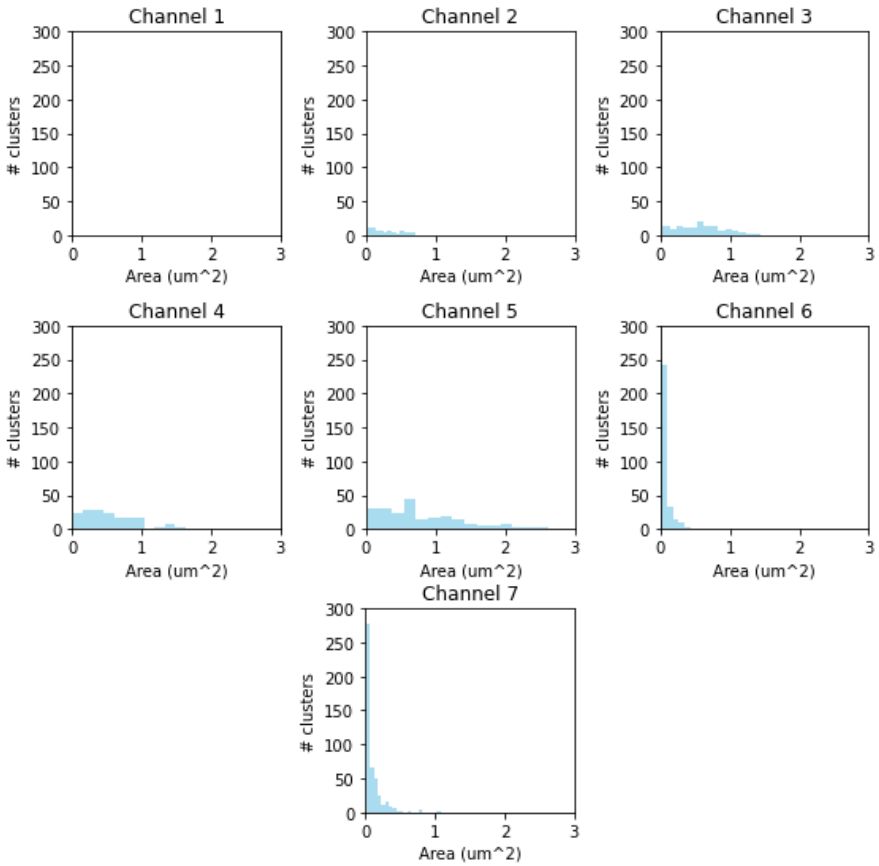


Figure 23 Histograms for cluster size vs cluster count for the second replicate. The histograms illustrate cluster counting vs. size for each receptor density concentration, enumerated as flow channels, in replicate 1. Channels 1 to 7 correspond to receptor density concentrations of Ch1 = 0 μm^{-2}, Ch2 = 3000 μm^{-2}, Ch3=10000 μm^{-2}, Ch4=20000 μm^{-2}, Ch5=50000 μm^{-2}, Ch6=150000 0 μm^{-2}, and Ch7=300000 μm^{-2} respectively. Channels 5, 6, and 7 exhibit the largest counts of clusters, with Channel 7, representing the highest count. Channel 5 shows the largest clusters.

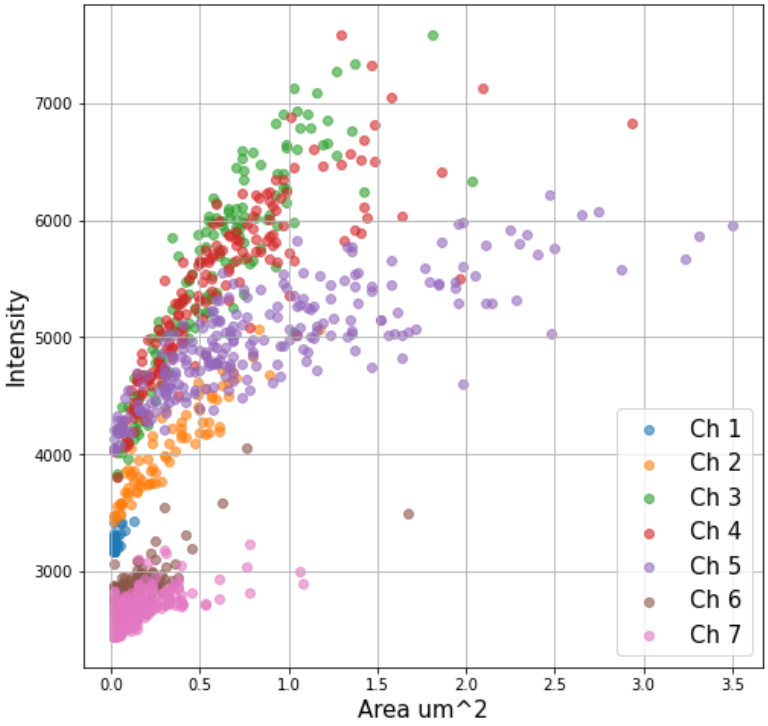


Figure 24 Scatter plot cluster size vs Intensity for second replicate. This scatter plot illustrates the size distribution of clusters per channel relative to the intensity registered during TIRF imaging. In contrast to the observations in the first replicate, Ch 1 and Ch 2 exhibit smaller low intense clusters. Conversely, Ch 3, Ch 4 and Ch 5 display the highest intensity values, with Ch 5 hosting the largest clusters registered. Furthermore, there is a noticeable reduction in intensity values for Ch 6.

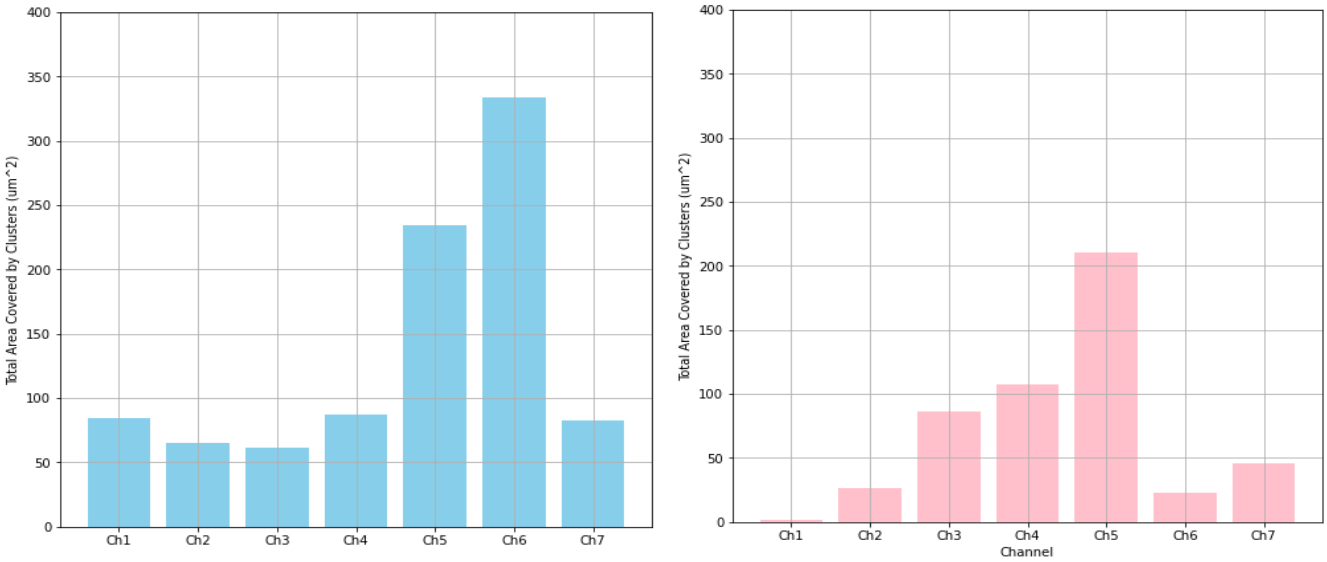


Figure 25 Total area covered by clusters for Design C. This image depicts two histograms representing the total area covered by clusters per channel in left: replica 1 (blue), right: replica 2 (pink). In replica 1, Ch5 and Ch6 exhibit the highest coverage, whereas in replica 2, Ch4 and Ch5 show the most coverage. Interestingly, Ch7 consistently exhibits one of the lowest coverages in both replicas.

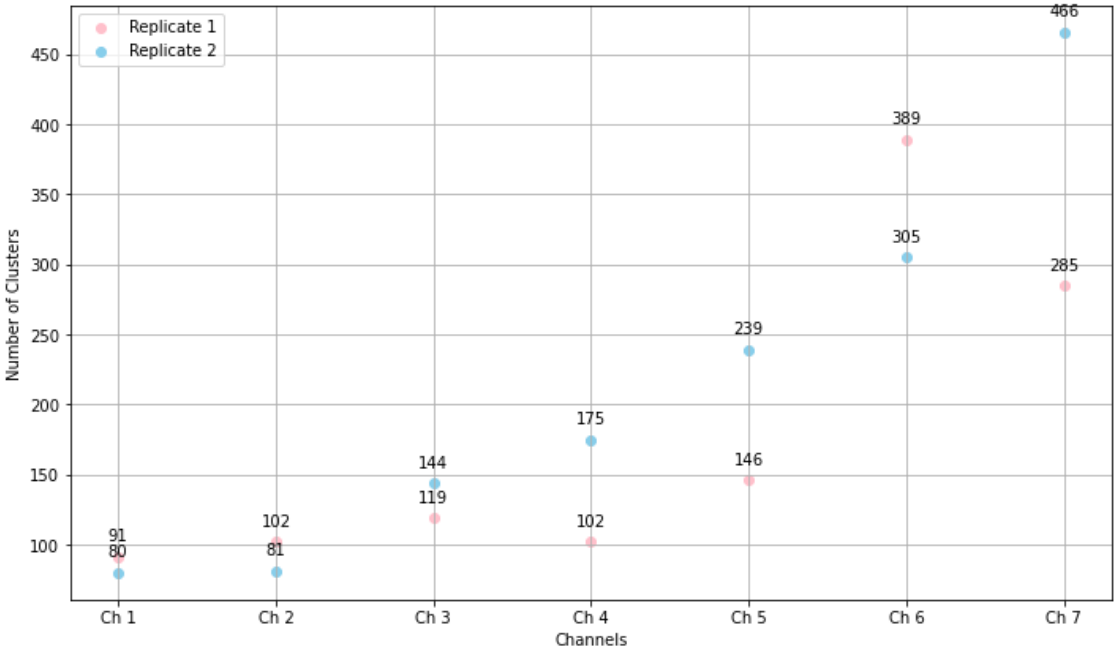


Figure 26 Comparative for cluster number vs channel for both replicates for Design C. This graph displays a scatter plot of the number of clusters vs the channel for both replicates of Design C, with replica 1 in blue, and replica 2 in pink. Ch6 consistently holds the highest cluster count for both replicates.

Finally, to facilitate comparison between the two replicates, we created two graphs depicting the total area covered by clusters and the comparative number of clusters for each, as shown in Figure 25 and Figure 26. Our observations revealed varying patterns of cluster formation and intensity across different receptor density levels (Ch 1 to Ch 7). In Figure 25, we noticed high cluster coverage, ranging from approximately 200 to 230 μm^2, for Ch 5 in both replicates. However, Figure 26 shows that Ch 6 has the highest count of clusters for both replicates, which coincides with the lowest bound fraction value registered in Figure 18. Interestingly, as receptor densities increased from Ch 3 to Ch 6, our scatter plots results indicate a reduction in cluster intensity values alongside an increase in cluster size and count, leaving Ch 6 to hold consistently the lowest intensities in clusters. This suggests a potential negative correlation between cluster formation and intensity values from Ch 3 to Ch 6.

Further measurements are required to confirm this relationship and to analyze the effects of cluster formation on the interpretation of adsorption profiles θ. While our results do not definitively establish a direct correlation between cluster coverage and average intensity value, they highlight the importance of considering cluster formation when calculating θ. Additionally, our findings indicate that the intensity of the clusters consistently exceeds background values, which should also be considered when analyzing average intensity values.

4 Discussion

4.1 Interpretation of experimental results

4.1.1 Validation of experimental model for superselectivity using Design A

The experimental validation of the protocol designed to study superselectivity in 6 arm/6bps DNA nanostars aimed to explore if we could replicate the outcomes for sigmoidal bound fraction θ, α, and binding constants K_A, K_{Intra}, and <n> obtained in the initial study by Linne (2022). Regarding the experimental bound fraction θ, our results demonstrated variation within expected ranges between the replicated results (DA) compared to the original (DA-CL), where even similar phenomena of reduction of coverage after reaching maximum adsorption (θ~1) and high coverage for low receptor densities were observed. This allowed us to demonstrate that these observed behaviors were not outliers and could indicate an unexplored phenomenon, as shown in Figure 12. Additionally, we performed a direct comparison of the experimental datasets representing the bound fraction for DA and DA-CL to establish if both datasets were statistically different, by using the non-parametric Mann-Whitney U p-value test. The results indicate no statistical difference between DA and DA-CL, suggesting they may represent the same phenomenon, supporting the reliability of the experimental protocol. To further validate the statistical results, the Mann-Whitney U test p-value between CL-DA and 6arm/4bps-CL was calculated as well, resulting in a value of 0.246 (p>0.05) (see Section 2.3.4). As this value suggests no significant difference between CL-DA and 6arm/4bps-CL, we interpreted it as an indication of how alike the data sets are, even though they are meant to represent different binding dynamics. However, we hypothesize that increasing both the number of experimental replicates and the resolution of receptor densities per design could enhance further differentiation of the evaluated curves. In summary, a qualitative and statistical assessment of the results was used to conclude the reliability of the experimental technique. We hypothesize the slight differences in slope value for θ can be resolved once more experimental replicates are done, and more receptor densities are explored.

4.1.1.1 Sensitivity analysis – results for Design A

To continue the assessment of reproducibility of the protocol using Design A (DA), the next step required calculating α, K_A, K_{intra} and <n>, from the adsorption profile from DA, and comparing to i) the reported values in (Linne, 2022), and ii) our reconstruction of the calculation using the reported data (DA-CL). The calculation of the variables comes from the theoretical definition of adsorption profile (or bound fraction) as described in Section 1.3.1, that later was adapted to quantify the change in the number of adsorbed nanostars on the target surface using intensity measurements, as shown in Section 2.3. To use θ, a least square fit model was used to generate θ_{fit}, with K_A and K_{intra} as fitting parameters. The calculation of α and <n> were derived from θ_{fit}.

Linne (2022) reports in her study in 6 arm DNA nanostars that for large receptor densities ($\sigma_R \sim 10^4$), and for low receptor densities ($\sigma_R \sim 10^3$), θ_{fit} shows a deviation from the data points, indicating the fit could not be describing θ accurately. Since our experimental results also showed the similar deviation at both high and low receptor densities, the sensitivity analysis aimed to explore the variation in fitting parameters K_A and K_{intra}, as well as the calculated α and <n>, as a result of varying the datapoints included in the fit. For the DA case, our observations show that the data points included in the fit influences considerably, even changing the interpretation of superselective behavior if the "wrong" set of points is included. For Case 1, all the points included the calculations of θ_{fit}, α and <n> described the sample as not superselective. For Case 2, once the low receptor density outlier was removed, DA calculations were superselective again. For Case 3, when both low receptor densities and high receptor densities outliers were

removed, there was no significant change with respect to Case 2. The latter allowed us to conclude for this case, the fit is more sensitive to low-receptor density outliers, a modification that should be considered for following experiments.

When quantitatively comparing the superselectivity parameters calculated for DA-CL and DA, irrespective of the sensitivity case chosen, there is an important variability among the values reported (Figure 16). This shows that the calculation of superselective parameters is a highly sensitive process, where small variations in slope values for θ_{fit} can lead to large variations in α, K_A and K_{intra}. We do not know if this variability can be entirely attributed to the difference in experimental replicates between DA (e.g. 2) and DA-CL (e.g. 3) and its associated uncertainty, or if it is also due to the use of a least-square fit to generate a non-linear function θ_{fit} that does not fully adjust to the theoretical sigmoidal behavior. Some other studies of selectivity on weak multivalent interactions also use a fit to calculate the adsorption profile from experimental data points, where the data points are fitted into a sigmoidal or exponential shaped function depending on the experimental observations (Scheepers et al., 2020). We point out the fit we used does not assume a shape on the fitted function, therefore this could cause the obtained α curves look less sharp and have lower α_{max} values, as well as to affect the obtained optimization parameters.

Referring to the effects of including or excluding data points for the fit when full adsorption was reached, it was noted that for Case 1, the calculations for DA-CL did coincide with the reported values (Linne, 2022) for the same set of data points. Nonetheless this case was also the one that showed the highest variability between DA and DA-CL, making it inadequate to conclude this is the best case for interpretation. Some studies done by Dubacheva and associates, show fits where the data points where full adsorption has been reached, are excluded from the fit function for θ (Dubacheva et al., 2015). The latter, together with our findings that the exclusion of these data points does change the α, <n>, K_A and K_{intra} values obtained, suggest that either it is required to develop a new analytical term θ capable of identifying when the sample has entered full bound regime, or to exclude the points from the fit calculation altogether.

Delving a little more into the possibility of holding an analytical definition of θ that accounts for not fully bound or fully bound states, we would like to discuss the equation chosen for our model to represent the bound fraction θ that depends on the definition of equilibrium avidity association constant K_A^{av}, as explained in Section 1.3.1. This variable describes the effective strength of multivalent interactions when more than one arm from the same DNA nanostar is bound. One can find similar definitions of K_A^{av} all over the literature, where some hold a term related to the number of accessible receptors, or free receptors, so that when the sample reaches full adsorption (e.g. free receptors ~ 0), θ and K_A^{av} should reflect this change (Curk et al., 2022) and possibly revert into strong binding definitions. In a way, we note that the definition of K_A^{av} used in our model does not have a mathematical definition that allows for changes in the number of receptors available for binding, as σ_R here refers to the total receptor density in certain area A. These could be compromising the interpretation, especially when $\theta \sim 1$ to later revert into lower coverage values, as observed in our results.

4.1.2 Evaluation of sequence mismatch on sticky end using Design A*

To analyze the influence of introducing changes in the binding affinity between ligands and receptors, we created a 2-nucleotide mismatch in the sticky end DNA complementary sequences between DNA nanostar's arms and receptors on the surface, referred to as Design A*. We thought this design would facilitate creating many weak binding receptors, as it is ideal for superselective systems (Martinez-Veracoechea & Frenkel, 2011), without having to introduce inert receptors, or increase the total number of receptors used. Nonetheless, our results show that introducing a possible reduction in free binding energy ΔG, resulted in the least superselective design when compared to DA-CL, DA and 6arm/4bps. The reported parameters α, K_A and K_{Intra} were the lowest of all the measured designs. Our interpretation is that the included mismatch interferes with the binding process in both initial bound and subsequent bounds, therefore a modification of this kind does not only affect the subsequent binding affinity K_{Intra}, but also K_A. The results also point out that the introduced reduction of binding affinity between ligands and receptors did not enhance superselective behavior, emphasizing the sensitivity of the proposed superselective system, and the delicate interplay of K_A and K_{Intra} to achieve higher superselective values.

This experiment highlighted DNA's unpredictable behavior within sensitive systems, such as the one employed in our study. Although the design aimed to introduce a two-nucleotide mismatch, with the remaining four nucleotides forming complementary base pairs, we hypothesize from the results an unexpected competition among potential binding configurations, as depicted in Figure 27. Studies show that imperfectly matched DNA sequences are still able to bind, and with important changes in the binding constants depending on the mismatches introduced (Cisse et al., 2012). According to the "Rule of Seven" in Watson-Crick base-pairing, at least seven contiguous base pairs are necessary for quick (microseconds) and efficient binding of DNA strands, which is defined as avoiding non-specific interactions. The same study also points out that mismatched DNA sequences required an increase in DNA concentration to observe binding, something we also found when trying to find the fully bound state (e.g. full adsorption in surface) with our DA* experiments. Therefore, this suggests that the presence of a two-nucleotide mismatch could have significantly disrupted the stability of the sticky end binding process, resulting in distinct configurations with highly different binding dynamics, something we can relate to the loss of superselective behavior we observed. Overall, the introduction of a two-nucleotide mismatch appears to create more than weak binding interactions, potentially resulting in nonspecific binding becoming the dominant outcome (Martinez-Veracoechea & Frenkel, 2011). This highlights the importance of considering the "Rule of Seven" when designing DNA sequences for superselective systems, something we can also extend to all the other designs.

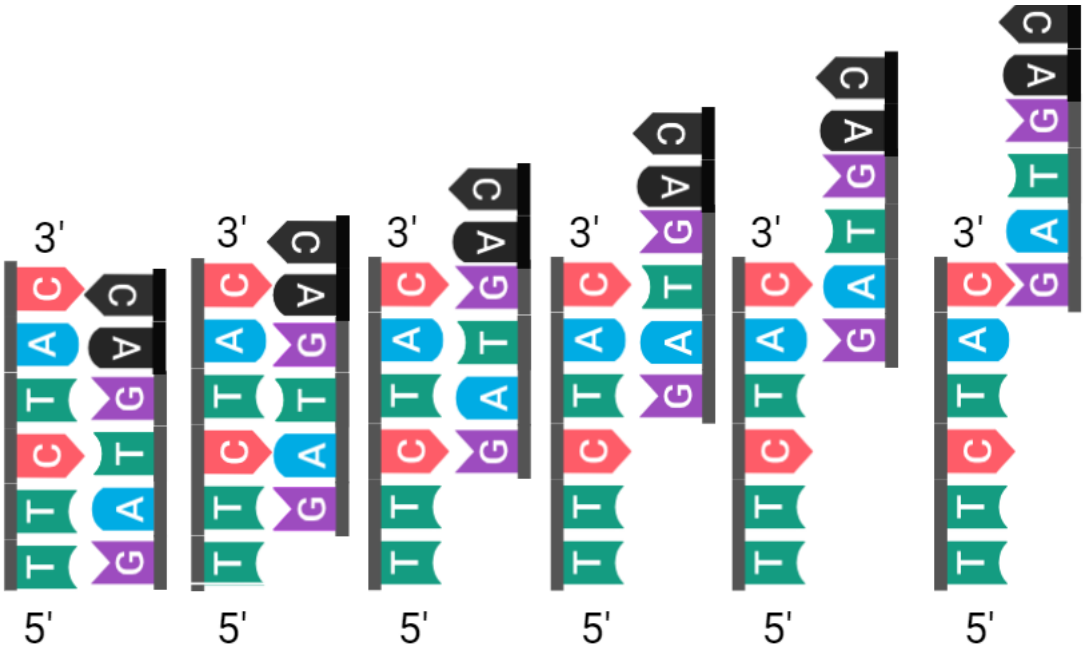


Figure 27 Possible configurations for unspecific DNA annealing of sticky ends for receptors and DNA nanostars after introducing a 2-nucleotide mismatch. We postulate that the "Rule of seven" for DNA annealing can give an explanation on why DA lost superselective behavior instead of gaining it. Figure created with BioRender.com*

*4.1.2.1 Sensitivity analysis -results for Design A**

The values obtained in Case 1 and Case 2 were the same, and there was a slight difference with Case 3; indicating that the results for this design are the least sensitive to variations in the data points for the fit calculations. Since this design did not show outliers for low receptor densities, in comparison to DA, we raise the question if this found stability depends solely on the binding kinetics of the sample, or it is also connected to the number of experimental replicates that were used for DA* (6 replicates). Nonetheless, we emphasize that the results from the analysis should not depend on the sample, for which more considerations should be made for the fit used.

4.1.3 Effects of increase flexibility on DNA nanostars - Design C

Regarding the effects of increasing the flexibility of the DNA nanostar structure by changing four Adenines into ten Thiamines in the vertex of the nanostar, it was found that these lead into unanticipated behavior in the measured bound fraction θ. The latter didn't show the expected sigmoidal function and even showed a negative slope, unlike seen for previous designs. We also observed a high variability among the replicates, unlike previous experiments. When exploring into the TIRF images obtained for this design, surprisingly, we noticed a significant number of clusters detected along certain receptor densities in both replicates, that look nothing alike to pervious results (See Figure 11, Figure 12, Figure 19, and Figure 22). This led our interest into analyzing if there was a preferred receptor density for the formation of clusters, and if there was a connection with the size and cluster number. Our results suggest the highest number of clusters formed in Ch 6 - σ_R = 150000 μm^{-2} (see Figure 20 and Figure 23), and the largest clusters formed in Ch 5 - σ_R = 50000 μm^{-2} (see Figure 24 and Figure 21), leading to the largest area coverage by clusters for these two densities (see Figure 25). Surprisingly, the particles seen later dissipate in size (for both replicates) and number (for only one replicate) when the maximum receptor density Ch 7 - σ_R =300000 μm^{-2} is reached. This finding suggests the existence of critical range of densities where the formation of these particles is thermodynamically preferred, but that unexpectedly doesn't correspond to the maximum possible receptor density, therefore we explored some other explanations.

As briefly described before, the formation of clusters on this experimental model would violate one of the main assumptions of no cooperation/interaction among guest particles (e.g. ligands), and the measurement of more than one DNA nanostar binding with one or more arms to the same receptors on surface (Linne, 2022) (see Section 1.3.1). Since this result would imply a reinterpretation of what the adsorption profile θ is measuring, we hypothesized that the observing cluster phenomenon could be affecting the final intensity reads obtained (see Figure 18). Surprisingly, our findings suggested a possible negative correlation between an increase in cluster formation and size, with a decrease in intensity values of the clusters for at least three receptor densities (σ_R = 10000 μm^{-2} to σ_R = 50000 μm^{-2} to σ_R = 150000 μm^{-2}). When comparing with the θ bound fraction for both replicates, it is observed that Ch 6 - σ_R = 150000 μm^{-2} holds consistently lower cluster intensities and lowest bound fraction values, which seems to support our hypothesis for negative correlation between average intensity measurements and cluster formation. Nonetheless, we underline that these finding may be somewhat limited by the number of replicates done, and the variability among them, for which further investigations are required to elucidate more about the role of cluster formation in superselectivity measurements.

Nonetheless, we would like to suggest a possible explanation for these results, as the formation of clusters among DNA nanostars has indeed been reported before. Biffi et. al. (2013) using 3- and 4-armed DNA nanostars with complementary sticky ends, concluded that for temperatures below T~42°C, all the DNA nanostars in solution on an electrolyte-rich buffer (NaCl in different concentrations for 3 or 4-armed) were found to undergo liquid-liquid phase separation into "coexisting small droplets". This indicates there is a phase separation between DNA nanostars in solution and those organized in droplets. The authors also observed that below this temperature, clusters grow larger as temperature reduces. Besides, they found that reducing the valence of these DNA nanostars leads to a reduction of the gas-liquid coexistence region, suggesting that the critical parameters of temperature and concentration for the coexisting region decrease as the valence is reduced. A similar critical temperature of T~42°C for DNA nanostars was also reported in independent studies using DNA nanostars as the system model for phase separation experiments (Hegde et al., 2023). Additionally, coarse-grain bead-spring model simulations for DNA nanostars with 3, 4 and 5 arms, support the previous results of phase change from unstructured fluid to a hydro gel phase for sufficiently low temperatures, although there is no direct mention of temperature values, but anything below allowing DNA to behave like gas (Naskar et al., 2023). This is an interesting point to mention, as for our experimental protocol, experiments are always done under 42°C and without strict temperature regulation, therefore increasing the chances of DNA phase separation in our samples.

Since the formation of clusters in DNA nanostars has been supported by recent studies, this leads us to consider the composition of these clusters. We suggest that phase separation could occur between i) DNA nanostars, ii) DNA receptors with DNA nanostars, and iii) DNA receptors alone. Although such phase separation wouldn't change

the overall concentrations of receptors or DNA nanostars in solution, it would influence the binding dynamics. This is critical because this affects the interpretation of bound fraction and superselectivity measurements.

Furthermore, the presence of phase separation challenges the assumptions of our theoretical models. Specifically, Equations 2 and 3, which describe the theoretical bound fraction Θ and its adaptation for experimental results, respectively, presume uniform receptor coverage. If phase separation occurs, this assumption is invalidated, leading to a scenario where receptors exist in a third state: unbound and unable to bind. In this scenario, the adsorption profile would most likely correspond to a mix of binding behaviors, including single DNA nanostars binding multivalently, and clusters of DNA nanostars that may not interact multivalently with surface receptors. Moreover, we propose this phenomenon could be responsible for the decrease in adsorption observed after reaching the first fully bound state in most of the experiments, as depicted in Figure 6 and Figure 12. Instead of indicating a reduction in receptor coverage, this decrease may actually reflect the clustering of nanostars and receptors, a phenomenon that could be misinterpreted as reduced adsorption.

Finally, we propose that the observed decrease in fluorescence intensity values with increasing cluster size suggests that clusters may not be bound to the surface. Instead, they appear to be deposited or condensed onto it. This hypothesis explains the measured contrast in TIRF microscopy intensities between the clusters and the background, as well as the overall reduction in average intensity values when cluster formation is promoted. While this potential phase separation phenomenon was more evident in Design C, likely due to its increased size and flexibility promoting larger clustering, our findings suggest that this phenomenon may be relevant to all the designs studied. Although we did not quantify these structures for Design A and Design A* in this thesis, TIRF images indicate the presence of cluster-like particles, albeit in smaller numbers and sizes compared to Design C (refer to Figure 11, Figure 15, Figure 19, and Figure 22). Therefore, we recommend further research to explore how this phenomenon could offer insights into reinterpreting the results obtained for superselective behavior with this experimental model.

5 Conclusion

Through a comprehensive investigation into the reproducibility of the experimental model using DNA nanostars to study superselective events, we examined the effects of altering the binding strength between DNA nanostars and receptors, as well as the influence of flexibility in the DNA nanostar structure for superselective behavior. This section summarizes the most relevant points for the reader.

Experimental Validation (Design A): Our analysis of experimental validation, focusing on Design A, revealed variation within expected ranges between replicated results and the original study by Linne (2022). Despite minor differences, statistical tests indicated no significant divergence, supporting the reliability of the experimental protocol. However, sensitivity analysis on the superselectivity calculations from the experimental data highlighted the influence of certain data points on the interpretation of superselective behavior for this design, emphasizing the need for careful consideration of the fit used for data analysis. Nonetheless, we acknowledge that this discrepancy could emphasize the need for further experimentation to enhance differentiation of evaluated bound fraction θ. Additionally, we suggest that further exploration into the analytical definition of θ and the equilibrium avidity association constant K_A^{av} could enhance our understanding of superselectivity dynamics in DNA nanostars, as a possibility to portray the so-called reduction of bound fraction at high receptor densities.

Evaluation of Sequence Mismatch on Sticky End (Design A*): The introduction of a 2-nucleotide mismatch in the sticky end DNA complementary sequences, referred to as Design A*, aimed to evaluate the impact of changes in binding affinity between ligands and receptors. Despite expectations that this design would promote weak binding receptors ideal for superselective systems, the results showed that Design A* exhibited the least superselective behavior compared to other designs. Parameters such as α, K_A, and K_{intra} were notably lower, indicating interference with the binding process and subsequent reductions in binding affinity. This suggests that modifications introducing a reduction in binding affinity using DNA mismatches do not necessarily enhance superselective

behavior, emphasizing the delicate interplay of K_A and K_{intra} in achieving higher superselective values on this system.

The experiment also revealed DNA's unpredictable behavior in sensitive systems, like the one used here. Despite the intention to create a specific mismatch, competition among stable and unstable binding configurations could have led to a reduction in specificity and selectivity among ligands and receptors. The analysis suggests that even a two-nucleotide mismatch can create more than weak binding interactions, potentially leading to nonspecific binding as the dominant case along the sample. For the future, we postulate considering the "Rule of Seven" for next DNA nanostar designs, to assure the complementary sequence for the sticky end is binding as expected, reducing the chances for non-specific interactions among DNA base pairs.

Sensitivity analysis results indicated stability in Design A* compared to other designs, raising questions about the factors influencing this stability, such as binding kinetics and the number of experimental replicates. However, it underscores the importance of ensuring that analysis results are independent of sample variations and that further considerations should be made for the fit used in future experiments.

Effects of Increased Flexibility on DNA Nanostars (Design C): The investigation into the effects of increasing flexibility on DNA nanostars, designated as Design C, has revealed unexpected behaviors in the measured bound fraction θ and the formation of clusters. Unlike previous designs, Design C displayed a negative slope in the bound fraction as the receptor density increased, indicating a clear deviation from the expected theoretical behavior. The observation of clusters along certain receptor densities suggests a preferential formation at $\boldsymbol{\sigma_R = 150000\ \mu m^{-2}}$, contrary to the assumption of no cooperation/interaction among DNA nanostars and in between receptors that was postulated for our experimental model.

We hypothesize that the unexpected behavior observed in Design C may be attributed to phase separation phenomena reported in previous studies, where DNA nanostars undergo liquid-liquid phase separation into coexisting small droplets at temperatures below 42°C to 45°C, which was never employed or regulated in our experiments. Since there appears to be a negative correlation between an increase in cluster formation and measured average intensity value, we postulate this could give potential explanation for the "reduction" of bound fraction after the sample reached full adsorption, as seen in previous studies and other designs used in this thesis. This phenomenon raises questions about the composition and implications of these clusters and suggests a need for reconsideration of the theoretical equations employed, and the interpretation on binding dynamics extracted from the model.

In conclusion, the findings from Design C highlight the complexity of DNA nanostar behavior and the need for comprehensive investigation into the factors influencing its superselectivity measurements. Further research is recommended to explore the potential implications of cluster formation in superselective systems using DNA nanostars, and to extend the analysis to previous design configurations.

5.1 Potential refinements in experimental design for further research

In this section, we summarize the key findings and insights gained from our investigation into superselective behavior using DNA nanostars. Building upon these findings, we offer recommendations for future research and potential refinements in experimental design to enhance the reliability and accuracy of our results.

- For our TIRF-experiments, we did not create a specific calibration profile for our samples, potentially leading to variability in intensity values measured. Although data normalization was applied to show general behavior with respect to coverage, Design C exhibited discrepancies that made us reflect in the importance of calibrating the microscope for our samples specifically. Therefore, performing a specific calibration profile for these samples could address this issue and reduce variability.

- Through discussion with TIRF experts revealed that slight changes in the thickness of the sample coverslip and glass slide could have significantly impacted intensity values, potentially affecting readings of superselectivity. Current coverslips and glass slides used in these experiments lack the recommended 1.5H thickness, possibly leading to skewed measurements. For future experiments, we recommend utilizing fine 1.5H glassware to ensure consistency and accuracy in TIRF measurements.

- Considering strict temperature regulation (at least making sure the DNA nanostars are always above the critical temperature 42°C to 45°C) in these experiments could provide more control on DNA nanostars' matter phases and structural conformations, leading to more understanding of what our experiments are showing. The majority of experiments (Lee et al., 2021) using this DNA nanostars are more focused on the structural changes and material properties of DNA condensing phases (Dubacheva et al., 2023) than on superselective behavior.

- Exploring strategies to enhance ligand-receptor interactions and reduce ligand cooperativity could improve superselective behavior. We know now that reducing the binding affinity by nucleotide mismatch is not a way to do so. Additionally, we recommend making sure that the sticky ends modification always respects the "Rule of Seven" to avoid non-specific interactions in the sticky end DNA strands.

- Recent studies emphasize the importance of spatial distribution of ligands and receptors in superselective behavior (Bila et al., 2022). Investigating different spatial configurations of receptors and ligands, along with multivalent pattern recognition, could provide valuable insights, as our experimental protocol does not regulate receptor displacement along the surface.

- Our results allowed us to reflect on how certain chemical modifications can also be disrupting the expected Watson-Crick interaction rules and may need to be taken more into account in the interpretation of superselective behavior for DNA nanostars. The inclusion of chemical modifications such as specific buffer-salt conditions or the use of fluorophores has been shown to lead to discrepancies between predicted and measured binding energies in DNA. Moreover, fluorophores have been found to alter the steady state and hinder the desired dynamic response of single-stranded DNAs (ssDNA) with Cy-dyes more likely to inhibit DNA binding dynamics compared to Atto-dyes (Idili et al., 2017). Acknowledging these differences is vital for accurately interpreting the binding dynamics of DNA nanostars. Additionally, the potential decay of fluorophores in current DNA nanostar samples should be taken into account, particularly for fluorophores older than three years.

- Likewise, fluorophore modifications have been reported to decrease the accessibility of ssDNA for base pairing and affect the dissociation constant (Jahnke et al., 2021). Therefore, it is then essential to consider that the binding affinity K_A of a DNA nanostar binding with a fluorophore arm differs from that of a non-fluorophore case. This variation in K_A influences the subsequent K_{intra} value, consequently affecting the overall dynamics and probability of development of superselective behavior. Taking these findings into

account could prove crucial when drawing conclusions about the general dynamics of superselective systems based on this model, as for our experiments we were assuming there is no binding energy cost for the arm that holds the fluorophore.

6 Bibliography

Antignani, A., Ho, E. C. H., Bilotta, M. T., Qiu, R., Sarnvosky, R., & FitzGerald, D. J. (2020). Targeting Receptors on Cancer Cells with Protein Toxins. *Biomolecules*, *10*(9), 1331. https://doi.org/10.3390/biom10091331

Arsiwala, A., Castro, A., Frey, S., Stathos, M., & Kane, R. S. (2019). Designing Multivalent Ligands to Control Biological Interactions: From Vaccines and Cellular Effectors to Targeted Drug Delivery. *Chemistry, an Asian Journal*, *14*(2), 244–255. https://doi.org/10.1002/asia.201801677

Berger, R. M. L., Weck, J. M., Kempe, S. M., Hill, O., Liedl, T., Rädler, J. O., Monzel, C., & Heuer-Jungemann, A. (2021). Nanoscale FasL Organization on DNA Origami to Decipher Apoptosis Signal Activation in Cells. *Small (Weinheim an Der Bergstrasse, Germany)*, *17*(26), e2101678. https://doi.org/10.1002/smll.202101678

Biffi, S., Cerbino, R., Bomboi, F., Paraboschi, E. M., Asselta, R., Sciortino, F., & Bellini, T. (2013). Phase behavior and critical activated dynamics of limited-valence DNA nanostars. *Proceedings of the National Academy of Sciences*, *110*(39), 15633–15637. https://doi.org/10.1073/pnas.1304632110

Bila, H., Paloja, K., Caroprese, V., Kononenko, A., & Bastings, M. M. C. (2022). Multivalent Pattern Recognition through Control of Nano-Spacing in Low-Valency Super-Selective Materials. *Journal of the American Chemical Society*, *144*(47), 21576–21586. https://doi.org/10.1021/jacs.2c08529

Capsid Constructors. (n.d.). Retrieved January 4, 2024, from https://capsidconstructors.github.io/lab-book/imaging/tirf.html

Carlson, C. B., Mowery, P., Owen, R. M., Dykhuizen, E. C., & Kiessling, L. L. (2007). Selective tumor cell targeting using low-affinity, multivalent interactions. *ACS Chemical Biology*, *2*(2), 119–127. https://doi.org/10.1021/cb6003788

Cisse, I. I., Kim, H., & Ha, T. (2012). A rule of seven in Watson-Crick base-pairing of mismatched sequences. *Nature Structural & Molecular Biology*, *19*(6), 623–627. https://doi.org/10.1038/nsmb.2294

Curk, T., Dobnikar, J., & Frenkel, D. (2018). Design Principles for Super Selectivity using Multivalent Interactions. In *Multivalency* (pp. 75–101). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781119143505.ch3

Curk, T., Dubacheva, G. V., Brisson, A. R., & Richter, R. P. (2022). Controlling Superselectivity of Multivalent Interactions with Cofactors and Competitors. *Journal of the American Chemical Society*, *144*(38), 17346–17350. https://doi.org/10.1021/jacs.2c06942

Dubacheva, G. V., Curk, T., Auzély-Velty, R., Frenkel, D., & Richter, R. P. (2015). Designing multivalent probes for tunable superselective targeting. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(18), 5579–5584. https://doi.org/10.1073/pnas.1500622112

Dubacheva, G. V., Curk, T., & Richter, R. P. (2023). Determinants of Superselectivity–Practical Concepts for Application in Biology and Medicine. *Accounts of Chemical Research*, *56*(7), 729–739. https://doi.org/10.1021/acs.accounts.2c00672

Frost, J. (2023, May 26). *Mann Whitney U Test Explained*. Statistics By Jim. https://statisticsbyjim.com/hypothesis-testing/mann-whitney-u-test/

Hegde, O., Li, T., Sharma, A., Borja, M., Jacobs, W. M., & Rogers, W. B. (2023). *Competition between self-assembly and phase separation governs high-temperature condensation of a DNA liquid* (arXiv:2301.06134; Version 2). arXiv. http://arxiv.org/abs/2301.06134

Heuer-Jungemann, A., El-Sagheer, A. H., Lackie, P. M., Brown, T., & Kanaras, A. G. (2016). Selective killing of cells triggered by their mRNA signature in the presence of smart nanoparticles. *Nanoscale*, *8*(38), 16857–16861. https://doi.org/10.1039/C6NR06154K

Idili, A., Ricci, F., & Vallée-Bélisle, A. (2017). Determining the folding and binding free energy of DNA-based nanodevices and nanoswitches using urea titration curves. *Nucleic Acids Research*, *45*(13), 7571–7580. https://doi.org/10.1093/nar/gkx498

Ijäs, H., Shen, B., Heuer-Jungemann, A., Keller, A., Kostiainen, M. A., Liedl, T., Ihalainen, J. A., & Linko, V. (2021). Unraveling the interaction between doxorubicin and DNA origami nanostructures for customizable chemotherapeutic drug release. *Nucleic Acids Research*, *49*(6), 3048–3062. https://doi.org/10.1093/nar/gkab097

Image Lab Software | Bio-Rad. (n.d.). Retrieved January 29, 2024, from https://www.bio-rad.com/en-nl/product/image-lab-software?ID=KRE6P5E8Z

Kapcan, E., Lake, B. P. M., & Rullo, A. F. (2023). Orchestrating Binding Interactions and the Emergence of Avidity Driven Therapeutics. *ACS Central Science*, *9*(4), 586–589. https://doi.org/10.1021/acscentsci.3c00242

Kiessling, L. L., Gestwicki, J. E., & Strong, L. E. (2000). Synthetic multivalent ligands in the exploration of cell-surface interactions. *Current Opinion in Chemical Biology*, *4*(6), 696–703. https://doi.org/10.1016/s1367-5931(00)00153-8

King, J. T., & Shakya, A. (2021). Phase separation of DNA: From past to present. *Biophysical Journal*, *120*(7), 1139–1149. https://doi.org/10.1016/j.bpj.2021.01.033

Lee, T., Do, S., Lee, J. G., Kim, D.-N., & Shin, Y. (2021). The flexibility-based modulation of DNA nanostar phase separation. *Nanoscale*, *13*(41), 17638–17647. https://doi.org/10.1039/D1NR03495B

Leroux, J.-C. (2017). Editorial: Drug Delivery: Too Much Complexity, Not Enough Reproducibility? *Angewandte Chemie (International Ed. in English)*, *56*(48), 15170–15171. https://doi.org/10.1002/anie.201709002

Li, W., Yang, X., He, L., Wang, K., Wang, Q., Huang, J., Liu, J., Wu, B., & Xu, C. (2016). Self-Assembled DNA Nanocentipede as Multivalent Drug Carrier for Targeted Delivery. *ACS Applied Materials & Interfaces*, *8*(39), 25733–25740. https://doi.org/10.1021/acsami.6b08210

Linne, C. (2022). *To bind or not to bind: DNA mediated multivalent interactions lead to superselectivity*. https://doi.org/10.4233/uuid:d2dbc4aa-02f5-43a1-a48c-b7ba1ad8ad67

Linne, C., Heemskerk, E., Zwanikken, J., Kraft, D. J., & Laan, L. (2023). *Optimality in superselective surface binding by multivalent DNA nanostars* (arXiv:2306.16885). arXiv. https://doi.org/10.48550/arXiv.2306.16885

Linne, C., Visco, D., Angioletti-Uberti, S., Laan, L., & Kraft, D. J. (2021). Direct visualization of superselective colloid-surface binding mediated by multivalent interactions. *Proceedings of the National Academy of Sciences*, *118*(36), e2106036118. https://doi.org/10.1073/pnas.2106036118

Lombardo, D., Kiselev, M. A., & Caccamo, M. T. (2019). Smart Nanoparticles for Drug Delivery Application: Development of Versatile Nanocarrier Platforms in Biotechnology and Nanomedicine. *Journal of Nanomaterials*, *2019*, e3702518. https://doi.org/10.1155/2019/3702518

Mahon, C. S., & Fulton, D. A. (2014). Mimicking nature with synthetic macromolecules capable of recognition. *Nature Chemistry*, *6*(8), 665–672. https://doi.org/10.1038/nchem.1994

Mammen, M., Choi, S.-K., & Whitesides, G. M. (1998). Polyvalent Interactions in Biological Systems: Implications for Design and Use of Multivalent Ligands and Inhibitors. *Angewandte Chemie International Edition*, *37*(20), 2754–2794. https://doi.org/10.1002/(SICI)1521-3773(19981102)37:20<2754::AID-ANIE2754>3.0.CO;2-3

Marin-Gonzalez, A., Vilhena, J. G., Perez, R., & Moreno-Herrero, F. (2021). A molecular view of DNA flexibility. *Quarterly Reviews of Biophysics*, *54*, e8. https://doi.org/10.1017/S0033583521000068

Martinez-Veracoechea, F. J., & Frenkel, D. (2011). Designing super selectivity in multivalent nano-particle binding. *Proceedings of the National Academy of Sciences*, *108*(27), 10963–10968. https://doi.org/10.1073/pnas.1105351108

Naskar, S., Bhatia, D., Lin, S.-T., & Maiti, P. K. (2023). Mechanistic insight into the structure, thermodynamics and dynamics of equilibrium gels of multi-armed DNA nanostars. *Physical Chemistry Chemical Physics*, *25*(11), 7847–7858. https://doi.org/10.1039/D2CP04683K

Ozalp, V. C., Eyidogan, F., & Oktem, H. A. (2011). Aptamer-Gated Nanoparticles for Smart Drug Delivery. *Pharmaceuticals*, *4*(8), Article 8. https://doi.org/10.3390/ph4081137

Saran, R., Wang, Y., & Li, I. T. S. (2020). Mechanical Flexibility of DNA: A Quintessential Tool for DNA Nanotechnology. *Sensors*, *20*(24), Article 24. https://doi.org/10.3390/s20247019

Scheepers, M. R. W., van IJzendoorn, L. J., & Prins, M. W. J. (2020). Multivalent weak interactions enhance selectivity of interparticle binding. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(37), 22690–22697. https://doi.org/10.1073/pnas.2003968117

Soliman, G. M., Sharma, A., Maysinger, D., & Kakkar, A. (2011). Dendrimers and miktoarm polymers based multivalent nanocarriers for efficient and targeted drug delivery. *Chemical Communications*, *47*(34), 9572–9587. https://doi.org/10.1039/C1CC11981H

Tjandra, K. C., & Thordarson, P. (2019). Multivalency in Drug Delivery-When Is It Too Much of a Good Thing? *Bioconjugate Chemistry*, *30*(3), 503–514. https://doi.org/10.1021/acs.bioconjchem.8b00804

Tong, G. J., Hsiao, S. C., Carrico, Z. M., & Francis, M. B. (2009). Viral Capsid DNA Aptamer Conjugates as Multivalent Cell Targeting Vehicles. *Journal of the American Chemical Society*, *131*(31), 11174–11178. https://doi.org/10.1021/ja903857f

Wang, J., Min, J., Eghtesadi, S. A., Kane, R. S., & Chilkoti, A. (2020). Quantitative Study of the Interaction of Multivalent Ligand-Modified Nanoparticles with Breast Cancer Cells with Tunable Receptor Density. *ACS Nano*, *14*(1), 372–383. https://doi.org/10.1021/acsnano.9b05689

Woythe, L., Tito, N. B., & Albertazzi, L. (2021). A quantitative view on multivalent nanomedicine targeting. *Advanced Drug Delivery Reviews*, *169*, 1–21. https://doi.org/10.1016/j.addr.2020.11.010

Zhang, Q., Jiang, Q., Li, N., Dai, L., Liu, Q., Song, L., Wang, J., Li, Y., Tian, J., Ding, B., & Du, Y. (2014). DNA origami as an in vivo drug delivery vehicle for cancer therapy. *ACS Nano*, *8*(7), 6633–6643. https://doi.org/10.1021/nn502058j

7 Appendices

7.1 Experimental protocols

Location: W:\staff-bulk\tnw\BN\LL\Shared\Valentina\Protocols

DNA Nanostar and Receptor Hybridization

1. Thaw the main stock (100 uM) of the desired DNA strands. In total you will need 6 different DNA stocks to obtain a DNA nanostar with 6 arms.
2. To obtain a 10 uM final DNA nanostar (NS) concentration, first, add 12 uL of TAE-MgCl buffer to an epje. Then, add 3uL of each main stock. In total you should have a final volume of 30 uL with a concentration of 100 uM.
3. At the same time, add 24 uL of TAE-MgCl buffer to an epje. Then add 3uL of each main stock for the receptor backbone and receptor.
4. Mix the epjes and spin them down with a small table centrifuge.
5. Place the epjes in a thermocycler and run the CL_HYBRID_NS program in the upstairs kitchen (left machine).
6. The program typically runs overnight and can be started around 16:00/17:00 and will be done on the next day in the morning.
7. When doing the insertion of DNA NS in the channels, in a 5mL Eppendorf dilute 2 uL of DNA-NS in 10 uM concentration into 1998 uL of TAE-MgCl buffer, obtaining a final concentration of 0.01 uM in each channel.

Small Unilamellar Vesicles of DOPC

Material:

- Chloroform, isopropanol, deionized water
- Three glass beakers
- DOPC
- TAE-NaCl buffer
- Gas tight syringes (non-fluorescent)

If you start with the dry aliquots from the freezer, you can start in step 6

1. Clean the gas tight syringe with Chloroform by pushing the liquid through the syringe at least 10 times.
2. Add 30 µL DOPC to a glass vial (there are already aliquots dried in the freezer)
3. Speed up the evaporation of the chloroform by gently blowing into the glass bottle with nitrogen
4. Place the glass bottle in a vacuum desiccator for at least one hour
5. Clean the gas tight syringe with Chloroform by pushing the liquid through the syringe for at least 10 times.
6. **X** Suspend the dried lipids in 375 µL of TAE-NaCl buffer and vortex it for 30 minutes (vortexer with Styrofoam on the top, maximum speed at 2).
7. Use the extruder with membrane size 0.05 um. Clean the syringes with isopropanol and water (deionized) by pushing the liquid through the syringe for at least 10 times.
8. Assemble the extruder and first test if the extruder is gastight by extruding deionized water. If that works fine, fill the syringes with 2x dilutes lipid solution. This means add 375 µl of TAE-NaCl buffer again.

It is normal to lose about 20% of the solution during the extrusion. For more information about the extrusion look here: https://www.youtube.com/watch?v=8ei62ZmLgSM&ab_channel=AvantiPolarLipids

9. Store the final solution in an epje in the fridge (maximum 5 days).
10. Clean the extruder parts and syringes thoroughly with first isopropanol and then rinse it with deionized water and let it dry before putting it back into the packaging.

DNA Nanostar binding experiments – sample preparation: DNA nanostar hybridization

Important: Hybridization always during the night before the experiment, do not store, prepare fresh for each day of experiment.

- Thaw the main stock (100 uM) of the desired DNA strands. In total you will need 6 different DNA stocks to obtain a DNA nanostar with 6 arms.
- To obtain a 10 uM final DNA nanostar (NS) concentration, first, add 12 uL of TAE-MgCl buffer to an epje. Then, add 3uL of each main stock. In total you should have a final volume of 30 uL with a concentration of 100 uM.
- At the same time, add 24 uL of TAE-MgCl buffer to an epje. Then add 3uL of each main stock for the receptor backbone and receptor.
- Mix the epjes and spin them down with a small table centrifuge.
- Place the epjes in a thermocycler and run the CL_HYBRID_NS program in the upstairs kitchen (left machine).
- The program typically runs overnight and can be started around 16:00/17:00 and will be done on the next day in the morning.
- When doing the insertion of DNA NS in the channels, in a 5mL Eppendorf dilute 2 uL of DNA-NS in 10 uM concentration into 1998 uL of TAE-MgCl buffer, obtaining a final concentration of 0.01 uM.

DOPC SUVs: Flow channel functionalization with lipid membrane and receptors

Important: Do not use cleaned glassware stored in KOH longer than 5 days. The KOH damages the surface over time and influences the formation of the lipid bilayer.

- Dry a pre-cleaned objective slide (Cleaned in 2%Hellmanex, Acetone and KOH (1M))
- *Place a piece of parafilm on the objective slide and cut out 7 times 1 mm broad channels
- *Place a cleaned coverslip on top and melt the parafilm on a heating plate @ 125°C
- *Press lightly with a pair of tweezers on the coverslip while the parafilm melts
- CRITICAL STEP: in case of using the double-sided polyimide tape (50mm width), skip all steps with * and proceed to past the flow channel template without heat, pressing lightly with your hand until the piece is attached completely.
- For each flow channel:
    - Add 13 µL TAE-NaCl buffer
    - Add 13 µL four times diluted SUVs
    - Incubate for 45 minutes in a closed chamber with a wet tissue
    - Wash four times with 13 µL TAE-NaCl buffer
    - CRITICAL STEP: Premix in epjes the desired DNA concentrations for the 7 channels receptors chosen
    - Incubate for 45 minutes
    - Wash four times with 13 µL TAE-NaCl-MgCl buffer
- From step 1.7 DNA Hybridization: Add 13 uL of the desired DNA nanostar concentration and wait 20 minutes before imaging.

Imaging parameters

TIRF 488 15%
Exposure time 50ms
Gain 50
Ellipse 100nm

Tips

- Negative control: Channel 1 no added receptors (measurement used for background corrections)
- Positive control: Channel 7 add maximum number of receptors (usually 300000). Here, we expect the system to be saturated and will be used to normalise your system. Be aware, this normalization is a little hand wavy

because sometimes we don't observe a clear saturation at this receptor density. The saturation depends a lot on the valency and interaction strength.

Increasing valencies with increasing number of sticky ends but with always 6 arms, exact DNA stock can be found in Lab_inventory_Linne.xlsx

Six-valent: X_3_1, X_3_2, X_4_3, X_5_4, X_6_5, X_6_6

Here, you can choose the number and location of the sticky ends by choosing, which sequence you use. The idea is to always keep the number of arms constant but vary the sticky ends and with that the valency. The sequences do not contain a sticky end and can be used as a base for this approach.

DNA receptor design

6bp: CL_157 (LongLInkerBAckbone) + CL_1108

Appendix: Dilutions for receptor densities

To obtain the desired receptor density on each channel, as seen in the table below, we must dilute the initial receptor DNA sample #1 (10 uM) into 3 additional concentration stocks.
#1 Is the initial concentration of 30ul DNA hybridized the day before.
#2 From this, we want to obtain a stock of 5 uM. Add 11.53 ul of #1 to 11.53 ul of TAE MgCl NaCl buffer.
#3 From this, we want to obtain a stock concentration of 1 uM. Add 4.62 ul of #2 to 18.4 ul of TAE MgCl NaCl buffer.
#4 From this, we want to obtain a stock concentration of 0.1 uM. Add 2.31 ul of #3 to 20.7 TAE MgCl NaCl buffer.
After having 4 different stock concentrations, proceed to dilute with TAE MgCl NaCl buffer to obtain the concentrations per channel, highlighted in green **2 to 7** on the table below. The channel **1** is the blank.
In the case you are using higher concentrations than 300000 DNA/um^2, make sure to calculate the right stock concentration.

Desired Linker DNA/um^2	Volume receptor [uL]	Stock concentration receptor [M]	Volume MgCl [uL]
1000 **2**	3.19	1.00E-07 **#4**	19.9
3000	9.56	1.00E-07	13.5
10000 **3**	3.19	1.00E-06	19.9
20000 **4**	6.38	1.00E-06 **#3**	16.7
50000 **5**	3.19	5.00E-06	19.9
100000	6.38	5.00E-06	16.7
150000 **6**	9.56	5.00E-06 **#2**	13.5
200000	6.38	1.00E-05	16.7
250000	7.97	1.00E-05	15.1
300000 **7**	9.56	1.00E-05 **#1**	13.5
600000	18.4	1.00E-05	4.6
750000	23	1.00E-05	0
1140000	7	5.00E-05	16
3750000	23	5.00E-05	0

7.2 TIRF DNA Nanostar Visualization

1. TIRF Microscope
    a. Turn on both switches on top of the microscope.
    b. Turn on the computer (takes long)
    c. In the meantime, you can lower the objective and change it to 100X.
    d. Put oil on objective and proceed to put sample.
    e. Open Metamorph when the computer has finished initializing.
    f. With the manual control in Z, slowly move the stage until you see that the sample touches the oil. For previous samples, we found z = 1790 to be the value for surface visualization. Turn on the PFS button to hear auditive validation that the PFS has found the sample.
    g. Make sure to start in Ch7, as it is the one that holds the largest fluorescence intensity.
    h. To localize the DNA NS attached to the surface, we will use the fluorophores Cy3. This excites at 555 nm max and emits at 569nm max. For this, we use the bypass 568 LP, to only detect the emissions of these. Make sure to insert this bypass.
    i. Using Metamorph, turn on the yellow laser TIRF 561, 11%, with an ellipse of 100 nm.
    j. Press the MDA button to open the camera.
    k. Focus the image until the structures are clear.
    l. Turn off the laser and change the bypass to 525/50. This will make sure to obtain most of the fluorescence only from the Atto 488 fluorophores on the DNA NS attached to the receptors.
    m. Now, open blue laser TIRF 488. Focus on the image again, it shouldn't be very far from the z position you had before.
    n. Establish the ellipse in 100nm, 11% TIRF 488.
    o. Check on the default values for EM gain = 50 and exposure time = 50 ms CHECK.
    p. Localize in the savings folder, put date and Ch.
    q. MDA requires you to choose 5 random locations and take 10 pictures on each. Once you have located the surface, you can move around the sample.
    r. Click Acquire.
    s. Wait until the pictures are ready, and check the PFS is still working, meaning the pictures are focused.
    t. Now take 30 random pictures without MDA along the channel.
    u. Go down with the manual control, find the first channel border, and continue with Ch6.
    v. Once you are done, repeat the procedure to test if time changes measurements. So, repeat using MDA from Ch1 to Ch 7.
    w. Save folder in Instruments…
    x. Make sure all lasers are off.
    y. Lower the stage as much as possible.
    z. Take the sample out.
    aa. Clean objective with isopropanol and Nikon wipes.
    bb. Change 100X objective to 40X objective.
    cc. Turn off the computer.
    dd. Turn off the microscope.

7.3 K_A, K_{intra} error propagation calculations from general error propagation formula

The error propagation formulas for Ka and Kintra were calculated using the concept of relative error propagation through mathematical functions with multiplication and division. When we have two variables X and Y related by a function $Z = f(X,Y)$ and we know the errors $\Delta X, \Delta Y$ we can estimate the error ΔZ using the following formula:

$$\frac{\Delta Z}{Z} = \sqrt{\left(\frac{\Delta X}{X}\right)^2 + \left(\frac{\Delta Y}{Y}\right)^2} \quad (1)$$

Where ΔX and ΔY are the relative errors of X and Y respectively. Since for K_A and K_{intra} we have the following relationships $K_A = \frac{C}{K_{intra}}, K_{intra} = \frac{1}{D}$, we can define:

$$f_1(C, K_{intra}) = K_A = \frac{C}{K_{intra}}, \qquad f_2(D) = K_{intra} = \frac{1}{D} \quad (2)$$

Where we obtain the following expressions:

$$\Delta K_A = \sqrt{\left(\frac{\Delta C}{C}\right)^2 + \left(\frac{\Delta K_{intra}}{K_{intra}}\right)^2} * K_A \quad (3)$$

$$\Delta K_{intra} = \frac{\Delta D}{D} * K_{intra} \quad (4)$$

The values were calculated for each case and are shown in Table 4.

7.4 Analysis Codes

All codes can be found in https://github.com/vquirogaf/MEP-TU-Delft, all codes and data sets can be found in WebDrive W:\staff-bulk\tnw\BN\LL\Shared\Valentina\Data_Analysis\All_Analysis_Codes, and all microscopy and electrophoresis images can be found in WebDrive W:\staff-bulk\tnw\BN\LL\Shared\Valentina\Data.